DTIC DEFENSE TECHNICAL INFORMATION CENTER

Information for the Defense Community

DTIC® has determined on 06/ 11 /2010 that this Technical Document has the Distribution Statement checked below. The current distribution for this document can be found in the DTIC® Technical Report Database.

☒ **DISTRIBUTION STATEMENT A.** Approved for public release; distribution is unlimited.

☐ **© COPYRIGHTED**; U.S. Government or Federal Rights License. All other rights and uses except those permitted by copyright law are reserved by the copyright owner.

☐ **DISTRIBUTION STATEMENT B.** Distribution authorized to U.S. Government agencies only (fill in reason) (date of determination). Other requests for this document shall be referred to (insert controlling DoD office)

☐ **DISTRIBUTION STATEMENT C.** Distribution authorized to U.S. Government Agencies and their contractors (fill in reason) (date of determination). Other requests for this document shall be referred to (insert controlling DoD office)

☐ **DISTRIBUTION STATEMENT D.** Distribution authorized to the Department of Defense and U.S. DoD contractors only (fill in reason) (date of determination). Other requests shall be referred to (insert controlling DoD office).

☐ **DISTRIBUTION STATEMENT E.** Distribution authorized to DoD Components only (fill in reason) (date of determination). Other requests shall be referred to (insert controlling DoD office).

☐ **DISTRIBUTION STATEMENT F.** Further dissemination only as directed by (inserting controlling DoD office) (date of determination) or higher DoD authority.

Distribution Statement F is also used when a document does not contain a distribution statement and no distribution statement can be determined.

☐ **DISTRIBUTION STATEMENT X.** Distribution authorized to U.S. Government Agencies and private individuals or enterprises eligible to obtain export-controlled technical data in accordance with DoDD 5230.25; (date of determination). DoD Controlling Office is (insert controlling DoD office).

TACTICAL SITUATIONAL AWARENESS SYMPOSIUM

DETECT PROCESS ACT

EWAT

Proceedings for
The Second Annual Symposium and Exhibition on

Situational Awareness
in the Tactical Air Environment

June 3 & 4, 1997
Patuxent River, Maryland

TABLE OF CONTENTS

* - executive summary only

This page intentionally blank.

INTRODUCTION TO THE PROCEEDINGS FOR THE 2ND ANNUAL SYMPOSIUM ON SITUATIONAL AWARENESS IN THE TACTICAL AIR ENVIRONMENT

Background

The 2nd Annual Symposium on Situational Awareness (SA) in the Tactical Air Environment was held on 3 and 4 June 1997 at the Naval Air Warfare Center Aircraft Division, Patuxent River, Maryland. The symposium was sponsored by the Electronic Warfare Advanced Technology Program, Naval Air Systems Command (PMA-272). The symposium was coordinated and hosted by the SA Integrated Product Team (IPT) at Patuxent River; points of contact: Karen Garner at 301-342-9285 and Tom Assenmacher at 301-342-0026.

Purpose

The objective of the symposium was to provide program managers, system developers, and system users with a heightened appreciation for potential SA improvements in tactical aviation through sensor fusion, decision aiding, adaptive automation, and training methods.

The symposium provided a unique opportunity for the 165 registered participants to discuss how SA influences design; learn new ways to research SA in the tactical air environment; learn the latest developments in SA-related technologies; discuss SA with experts on panels and on a one-to-one basis; and network with a variety of SA researchers from government, industry, and academia.

Description of Proceedings

The SA IPT Summary is included in the proceedings to provide background on how the symposium started and the specific interests and ongoing work performed by the IPT at Patuxent River.

Twenty nine presentations were given during the symposium. This document contains formal papers based on those presentations. Where papers are not available, executive summaries previously printed in the symposium notebook are reprinted in the proceedings to provide comprehensive documentation of topics and authors for your reference.

There are several papers not related to symposium presentations, however, the topics are directly related to SA. We included these papers in the proceedings as a service to you and hope that you find this additional information useful.

Personal Note

I sincerely hope that those who participated in the symposium left with ideas and insight for improving SA, new contacts to help in your work, and renewed faith that there is indeed something we can do to improve the tactical viewpoint. For those reading these proceedings, I hope the document is a useful reference for SA information and contacts.

Karen T. Garner
Situational Awareness Integrated Product Team Leader
Code 4.6.4.2, Bldg. 2187, Suite 2280
Naval Air Warfare Center Aircraft Division, Patuxent River, MD

This page intentionally blank.

SITUATIONAL AWARENESS
INTEGRATED PRODUCT TEAM SUMMARY

Karen T. Garner
Naval Air Warfare Center
Aircraft Division
Patuxent River, MD

Thomas J. Assenmacher
ARINC, Inc.
California, MD

BACKGROUND

The Situational Awareness (SA) Integrated Product Team (IPT) is sponsored by the Electronic Warfare Advanced Technology (EWAT) Program. The team was formed in response to fleet tactical aircraft aviators ranking SA as a critical mission concern. The team developed a Navy definition of SA based on existing research and investigation of tactical aircraft electronic combat problems. A clear link between the threats encountered and crew-system integration with the various avionics/sensor systems is provided by the following example taken from the investigation:

> Current tactical EW systems lack integration and sufficient fusing/merging/ correlation of sensor data. They can only operate because the operator is in the loop, using multiple independent equipments, responding to threats keyed by radar warning receivers and visual sightings. The aircrew tends to be overworked, responding to threats at the wrong time, overusing limited expendables, and often giving no defensive response since some missiles are never detected.

Definitions of SA range from the global "staying ahead of the aircraft" or "knowing what's going on so you can figure out what to do" to the more comprehensive and formal definition now included in several Navy specifications:

> Operator SA is comprised of detecting information in the environment, processing the information with relevant knowledge to create a mental picture of the current situation, and acting on this picture to make a decision or explore further.

PURPOSE

The primary purpose of the SA IPT is to acquaint acquisition and program managers, system developers, and the fleet with SA achievement and the benefits of successful cockpit SA management strategies. Research on current initiatives in this evolving area of interest is being coupled with known problem areas in the electronic combat arena to improve existing systems and guide future system designs.

GOALS/OBJECTIVES

This is our second year coordinating and hosting the SA symposium. The symposium provides a forum for SA information exchange between academia, industry, tactical aircraft platform and program managers, and aircrew. This information exchange heightens individual perception and appreciation of SA in the tactical environment and will eventually lead to efficient avionics system design and employment.

We've compiled SA guidelines based on lessons learned and research of current and emerging technologies. The 1992 Patuxent River investigation of tactical aircraft electronic combat problems was the impetus for this effort and was referred to frequently in an effort to "answer" some of the interface issues

documented in that report. Research of the interface guidelines consisted of gleaning appropriate material from numerous Human Computer Interface documents.

We believe that SA can be improved by enhancing both the design of data-to-information processing and the display of this information to the crew. Applying the guidelines, along with crew-system integration and aircrew consultation during the development cycle, will result in an efficient and effective presentation of necessary information to the aircrew with high confidence that the latest thinking has been employed.
The guidelines are published as a Technical Memorandum (Report No. NAWCADPAX--96-268-TM) of 8 January 1997. Copies are available from the authors above upon request.

Our rapid prototyping capabilities expanded during the past year to include the Air Combat Crew Evaluation Simulator System (ACCESS). This reconfigurable system is used for research, prototyping, human factors engineering, and threat environment simulation. The pilot station contains multipurpose displays and controls, instruments, stick, throttle, other control panels and indicators. High resolution Silicon Graphics, Incorporated (SGI) monitors simulate glass aircraft displays (multipurpose display (MPD), radar warning receiver (RWR), digital data display, etc.). An active aircraft replica faceplate overlays the monitor providing the look and feel of an actual aircraft. Electromechanical replicas are used to replace gauges that cannot be displayed on the MPDs. The cockpit station has a headset for voice communication, aural tones and cues, and aircraft noise simulation. A head-up display (HUD), generated by an SGI computer, is mixed with the out-the-window (OTW) imagery to provide composite OTW and HUD image. The system also includes an evaluator station for monitoring, participating in, and analyzing performance; a computation system that drives the ACCESS; and sophisticated software to support weapons system employment, instrument flight and approaches, and emergency procedures. ACCESS software consists of the same components as in an actual aircraft. RWR, Infrared Search and Track, and Infrared warning receivers are modeled as well as imaging sensors such as Forward-Looking Infrared pods and LANTIRN. The navigation systems include Inertial Navigation System, Tactical Control and Navigation, VHF Omni Range, and Instrument Landing System. The system is also used for mission generation and editing. New scenarios may be developed, existing scenarios modified, or scenarios deleted. Rapid prototyping is a key technology that enables exploration of various presentation methods and allows selection of formats that result in the best system performance based on user evaluation.

TECHNICAL AREAS OF INTEREST

We feel that "human factors" as well as equipment enhancements should be considered in any attempt to improve SA. A human factors analysis is the first step toward improving SA, whether a new start program or a platform upgrade. First, determine the overall objectives and performance requirements of the system. Describe the operating conditions and, from these, develop scenarios that exercise the systems during typical missions, mission failures, and in multimission environments. Second, define the functions required to accomplish the mission. Any mission can be broken down into a sequential set of system functions whose completion represents the successful accomplishment of that mission. Third, derive the operator tasks and cross check them against the scenarios. For each task in each function, address the operator input into the system and the system output to the operator. To ensure aircrew performance payoff, new system design or system upgrade design should be undertaken with a human centered approach rather than being hardware driven.

SA improvement initiatives incorporating sensor fusion are ongoing throughout industry and the services. The challenge is not in collecting information, but in correlating and presenting it in a coherent, tactically useable format. Sensor fusion should provide a complete, timely assessment of situations and threats. It can meet this challenge through software management schemes, processing architectures, and correlation systems that automatically sort and merge available sensor and intelligence data. There must be a common time base and common spatial reference for sensor information. A confidence level should be assigned to each source based on its reliability and credibility. Information ambiguity should be resolved during analysis and integration. By taking a closer look at data provided by existing sensors, incorporating

advanced technology techniques, and integrating improved symbology, a more comprehensive, focused " picture" of the environment can be presented to the aircrew. Sensor fusion should not result in aircrew confusion.

The SA IPT is leading a three-year effort beginning in FY98 to improve platform lethality and survivability by decreasing threat classification ambiguities through aircraft sensor data fusion. During the first year, the team will prototype and demonstrate basic architecture (i.e., fusion engine, data conversion engine, response engine, and operator display and interface) with emphasis on existing EWAT products (e.g., sensor and response systems/concepts) and systems currently available. This will be followed by the actual execution of the design of the concept envisioned and the validation of the resulting technique/system using simulation and/or analysis. The demonstration phase in year 3 involves the integration of the resulting technique/system onboard the QF-4 platform to conduct live demonstrations. As new systems and EWAT products with applicability to sensor fusion come on-line, additional phases will be structured and executed to eventually realize the full capabilities allowed sensor fusion architecture. There are many technical challenges to be resolved and we are currently developing our approach to be ready to begin this effort as soon as FY98 funds become available.

Even with today's highly accurate and effective weapons, tactical airborne mission effectiveness depends on the aircrew achieving and maintaining a high level of SA throughout the entire mission. This can be done by first recognizing the capabilities and limitations of the human operator and designing/upgrading airborne systems based on these factors.

This page intentionally blank.

SENSOR FUSION

* - Executive Summary only

This page intentionally blank.

LESSONS LEARNED IN USING SITUATIONAL AWARENESS TO EVALUATE SYSTEMS

Valerie J. Gawron
Calspan SRL Corporation
Flight Research Department
Buffalo, New York

INTRODUCTION

Designers of sensor fusion, decision aiding, and adaptive automation systems have expected these systems to enhance Situational Awareness (SA) and hence mission effectiveness. These expectations have not always been met. The focus of this paper is to describe possible reasons for these results. These reasons are presented in the form of eleven lessons learned.

LESSON ONE

There may be dissociations among performance, workload, and SA. The dissociation between performance and workload may be partially explained by the inverted U relationship between them (Figure 1). Specifically, performance is optimum at moderate levels of workload and degraded at either low (sometimes called underload) or high (sometimes called overload) workload. The classic example is maintaining altitude during cruise (underload) and during landing (overload).

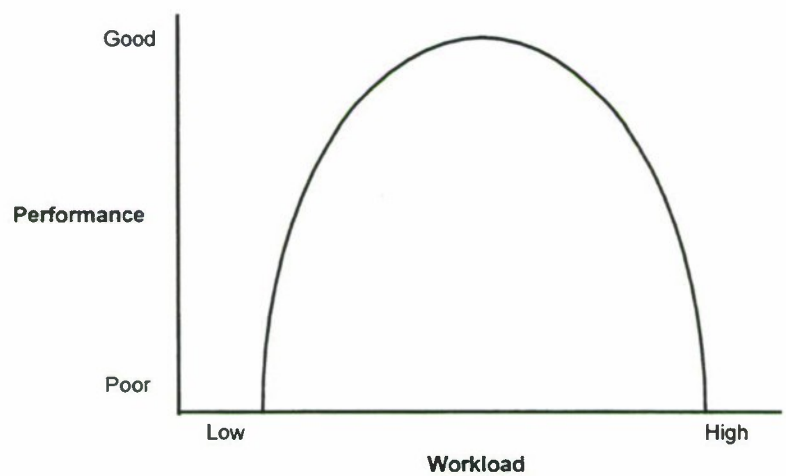


Figure 1. Relationship Between Performance and Workload

More research on this dissociation has been done at the University of Illinois. Specifically, Yeh and Wickens (1988) defined dissociation as a workload measurement error that occurs whenever there is a difference in implied workload between measures of primary task performance and measure(s) of subjective workload. For example, primary task performance degrades (implying increased workload) and secondary measures of workload decrease (implying decreased workload).

A multi-dimensional concept, workload is defined as the effort expended by the human operator in accomplishing the imposed task requirements (Table 1). The task requirements (taskload) are the goal to be achieved, the time permitted to perform the task, and the performance level to which the task is completed. The factors affecting the effort expended are the information and equipment provided, the task environment, the subject's skills and experience, the strategies adopted, and the emotional response to the situation. This definition provides a testable link between taskload and workload. For example (paraphrased from an example given by Azad Madni, Vice President, Perceptronics, Woodland Hills, CA,

on results of an Army study), the workload of a helicopter pilot in maintaining a constant hover may be 70 on a scale of 0 to 100. Given the task of maintaining a constant hover and targeting a tank may also be 70. The discrepancy results from the pilot self-imposing a strict performance requirement on hover-only (no horizontal or vertical movement), but relaxing the performance requirement on hover (a few feet movement) when targeting the tank to keep workload within a manageable level. These definitions allow taskload and workload to be explainable in real work situations. These definitions also allow logically correct analysis of taskload and workload. Realistically speaking, workload can *never* exceed 100% (a person cannot do the impossible). Any theories or reported results which allow workload to exceed 100% are not realistic. However, as defined, taskload may exceed 100%. An example is measuring "time required/time available." By the proposed definition, this taskload measurement may exceed 100% if the performance requirements are set too high (thereby increasing the time required) or the time available is set too low. The bottom line is workload cannot exceed 100%, even if taskload does exceed 100%.

TABLE 1

Factors in Workload Assessment

Workload Aspect	Representative Questions	Candidate Measures	Comments
Sustained Workload	What was the average overall workload?	Subjective rating scales	
	How will various intensities of sustained workload affect performance?	Overall performance on salient tasks	
Momentary Workload	What was the magnitude of workload during peak periods?	Subjective rating scales	Global measures of task performance are inappropriate
	How was human performance affected during periods of high demand?	Secondary task performance	
Reserve Capacity	What margin of full performance did this task require?	Secondary tasks	Primary task measures are inappropriate
	How much more can the operator handle effectively?	Subjective rating scales	

Yeh and Wickens (1988) identified five pairs of conditions in which the relationship between performance and subjective measures of workload imply different effects on pilot workload. In the first condition, termed motivation, performance improves and subjective ratings of workload increase. In the second condition, which Yeh and Wickens termed underload, as demand increases, performance remains the same but subjective measures of workload increase. In this condition, performance implies that workload remains constant while the subjective measures imply increased workload. In the third

condition, resource-limited tasks, as the amount of invested resources increases, performance degrades and subjective workload measures increase. However, the proportion of the change in subjective ratings is greater than the proportion of the change in performance. In this condition, performance implies that workload increases somewhat while subjective workload measures imply that workload increases greatly. In the fourth pair of conditions, comparison of dual-task configurations with different degrees of competition for common resources, as demand for common resources increases, performance degrades and subjective workload measures increase. This time, however, the change in performance is greater than the change in subjective ratings. In this condition, performance implies that workload increases greatly, while subjective workload measures suggest that workload increases slightly. In the fifth pair of conditions, which Yeh and Wickens termed overload, as demand increases, performance degrades while subjective measures remain unchanged. In this condition, performance implies that workload increases while the subjective measures imply that workload remains the same.

Yeh and Wickens defined these differences in implied workload as dissociation and they suggested that dissociation occurs because, in these five conditions, different factors determine performance and subjective workload measures. These factors are listed in Table 2. As can be seen from Table 2, there is only one factor common to both performance and subjective measures of workload: amount of invested resources. However, there are different types of resources. For example, Yeh and Wickens (1988) defined four types of resources: 1) perceptual/central versus response stage, 2) verbal versus spatial code, 3) visual versus auditory input modality, and 4) manual versus speech output modality.

TABLE 2

Determinants of Performance and Subjective Measures of Workload

Measure	Primary Determinant	Secondary Determinant
Single-Task Performance	Amount of Invested Resources	Task Difficulty Subject's Motivation Subjective Criteria of Optimal Performance
	Resource Efficiency	Task Difficulty Data Quality Practice
Dual-Task Performance	Amount of Invested Resources	Task Difficulty Subject's Motivation Subjective Criteria of Optimal Performance
	Resource Efficiency	Task Difficulty and/or Complexity Data Quality Practice
Subjective Workload	Amount of Invested Resources	Task Difficulty Subject's Motivation Subjective Criteria of Optimal Performance
	Demands on Working Memory	Amount of Time Sharing Between Tasks Amount of Information Held in Working Memory Demand on Perceptual and/or Central Processing Resources

(Adapted from Yeh and Wickens, 1988)

The importance of these dichotomies has been supported in many experiments. For example, Wickens and Liu (1988) reported greater tracking error when a one-dimensional compensatory tracking task was time shared with a spatial decision task (same code: spatial) than with a verbal decision task (different codes: spatial and verbal). Tracking error was also greater when the response to the time shared decision task was manual (same modality: manual) rather than verbal (different modality: manual versus speech).

Derrick (1985) analyzed performance on 18 computer-based tasks and a global subjective estimate for workload on each task. The tasks were performed in four configurations: 1) single easy, 2) single hard, 3) dual with the same task, and 4) dual with different tasks. He concluded: "If a task was increased in difficulty by adding to the perceptual and central processing resource load (and thus performance declined), people rated the workload as higher than for the same task with lower perceptual and central processing resource demands. However, tasks made more difficult by increasing the resource demands of responding (again worse performance) did not produce increased workload ratings" (p. 1022). In addition, "performance in the Dual Self configuration was worse than in the Dual Other configuration, but the workload ratings were essentially equivalent for these two conditions" (p. 1022).

Derrick summarized his findings as well as those of Yeh and Wickens (1984) in Table 3.

TABLE 3

A Theory of Dissociation

	Sources	Performance Decreases	Subjective Difficulty Increases
1	Increased single task difficulty	4	3
	Perceptual/cognitive	2	2
	Response	2	1
2	Concurrent task demand	3	4
	Same resources	2	2
	Different resources	1	2

The relationship between the factors which determine performance and subjective measures of workload and the resource dichotomies is complex. Greater resource supply improves performance of resource-limited single tasks but not of data-limited single tasks. For dual tasks, performance degrades when the tasks compete for common resources.

The importance of accurate workload measurement during system evaluation cannot be overstated. As Derrick (1985) stated: "A workload practitioner who relies on these [subjective] ratings [of workload] rather than performance will be biased to choose a nonoptional system that requires operators to perform just one task rather than the system that demands dual task performance" (p. 1023). "The practitioner who relies solely on subjective data may likely favor the system that has serious performance limitations, especially under emergency conditions" (p. 1024). "A system design option that increases control demands, and ultimately degrades performance under high workload conditions, will not be rated as a problem in the normal single task workload evaluations" (p. 1024). Therefore, the potential for inaccurate interpretation of the results is worrisome.

When dissociation occurs, Yeh and Wickens (1988) suggest that performance measures be used to provide a "direct index of the benefit of operating a system" (p. 118). Further, subjective measures should

be used to "indicate potential performance problems that could emerge if additional demands are imposed" (p. 118).

Yeh and Wickens (1988) suggested the following research: 1) devise difficult single-task and easy dual-task conditions that require equivalent resources to know whether subjective measures are more sensitive to the total amount of invested resources or more sensitive to cognitive demand on working memory, and 2) examine how affective variables such as preference, attitude, and volition influence subjective workload. Another line of research should be undertaken as well, specifically the development of a ratio scale measure of workload. Some work towards this end has been initiated at Calspan with the development of a time limit measure of workload.

The inverted U relationship may also occur between SA and workload (Figure 2). Specifically, optimum SA occurs at moderate workload. At too little workload, a person minimizes his or her sensory processing. A good example is vigilance decrement. If signals are few and far between (low workload), operators begin to minimize the number of displays monitored and/or the frequency of monitoring these displays (minimum sensory processing). At too high workload, a person focuses on only one stimulation. Pilots in a dog fight don't see the fuel warning flashing across the entire HUD. They only see the enemy aircraft they are trying to destroy.

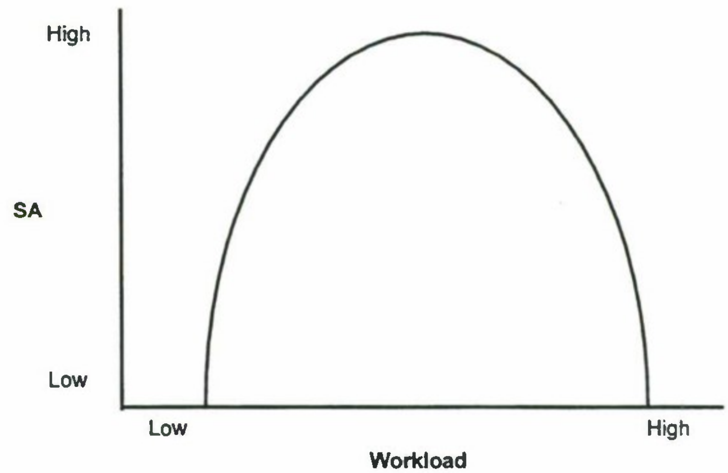


Figure 2. Relationship Between SA and Workload

There seems to be the same inverted U relationship between SA and performance. In a recent ground-based simulation of a threat warning device that used sensor fusion algorithms, pilots with less sensitive sensors (low SA) failed to evade threats (poor performance). Pilots with highly sensitive sensors (high SA) also failed to evade threats. They "died smart" but they did die.

LESSON TWO

Operators can have too much SA. In another sensor fusion display evaluation, one display presented all the relevant information on all the threats that could destroy the aircraft. Pilots' performance (measured as the distance penetrated past the first threat) for this full SA system was degraded compared to a system that showed only a portion of the threats. During the debriefs, pilots stated that "they knew too much" in the full SA system. They tried to use this knowledge to develop an optimum course. They failed. In the system that showed just a portion of the threats (Figure 3), they were "fat, dumb, and happy" as well as more successful.

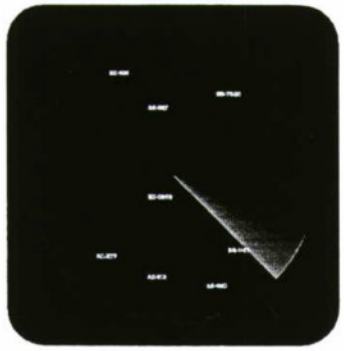


Figure 3. Minimalistic Sensor Fusion System

LESSON THREE

All three levels of SA must be measured during any system evaluation. The study of SA is in its infancy and, as such, suffers from a lack of a commonly accepted definition. For example, Fracker (1988) lists five definitions of SA. Each of these definitions, however, contains the same components: 1) the internal states of the system, i.e., what all the subsystems are doing; 2) the external states of the system, i.e., the integrity and functioning of the components of the system boundary; 3) the relationship between the system and its environment, i.e., its status relative to pertinent external environmental variables; and 4) the environment in which the system exists, i.e., objects and conditions external to the system that may affect the efficient and safe operation of the system. These components are illustrated in Figure 4. The painter knows what all his subsystems are doing, i.e., "My hand held the paint can." He also knows the state of the bench, i.e., "The bench is painted" and the environment, i.e., "Paint is wet." What he lacks is knowledge of the bench's relationship with the environment, i.e., "Wet paint comes off on clothes."

Endsley (1991b) used these four components to define three levels of SA: Level 1, perception of the elements in the environment; level 2, comprehension of the current situation; and level 3, projection of future status. All three levels of SA must be measured during any system evaluation.

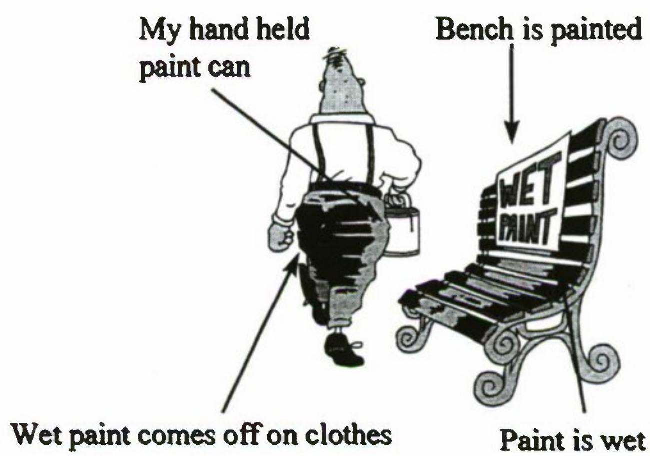


Figure 4. Elements of Situational Awareness

A decision aiding system automatically detected and identified all entities. The human operator was expected to project the future actions of the entities. Operators did worse with the system than without it. Extensive debriefing revealed that the operators felt they did not have the opportunity to assess the behavior of the entity during the detection and identification. Therefore, they could not project its future actions since they had no knowledge of its past actions.

LESSON FOUR

There are individual differences. This is nothing new. Every pilot knows who has the best SA in the squadron. Ironically the designer expects the best pilot to show the *greatest* increase in SA due to a new SA-enhancement system. But the best doesn't have much room for improvement. However, the worst does. One of the most successful SA-enhancement systems designed brought the lowest SA pilot at par with the best. Much to the surprise of everyone in the squadron. Individual differences also exist in pilots with equally good SA skills. Some maintain SA best with spatial information, others with auditory or text or symbology. Another successful SA enhancement system recognized these differences and allowed pilots to customize the display.

LESSON FIVE

The right intentions but a bad design still results in bad SA. An adaptive system perfectly compensated for decrements in pilot performance (fuel management). But the system informed the pilot of the actions it had taken (e.g., move fuel to another tank) using a low contrast display that washed out in bright sunlight. During an engine fire, the pilot did not eject because he "knew" the fuel tanks on the fire side were empty. The resulting blast destroyed the aircraft.

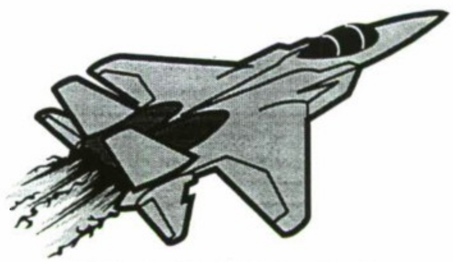


Figure 5. Fuel Management

LESSON SIX

Responsibility is critical to assessing SA. In a recent exercise, a team of six soldiers were instructed to retake a building seized by "terrorists." The team mapped out roles and responsibilities for each member but did not rehearse beforehand. No one was specifically told to monitor the data collection team. The members of this team had been through many exercises but never as a team and ironically none of them had ever been given the responsibility for monitoring the data collection team. Also ironically all members of the data collection team had participated in many previous exercises but never as a team. When a freak power outage crashed the data collection computers, the data collectors physically moved into the line of fire between soldiers and "terrorists." Luckily injuries and damage were minor. Afterwards every member of the soldier and data team stated "It wasn't my job to watch out for them." Persons who do not think they are responsible for being aware of an entity, do not try to maintain SA on that entity.

LESSON SEVEN

Rules are made to be broken. A better mousetrap will trap more mice right? A better threat identification system will identify more threats and give operators better SA right? Not necessarily. Fusing sensor images usually makes identification of threats easier. But in one system, operators felt that the work of manually fusing the information enhanced their SA of these entities as well as their SA of the status of the sensors.

Figure 6. A Better Mousetrap

LESSON EIGHT

Words are critical in rating SA.

Situational Awareness Rating Technique. *General description* - An example of a subjective measure of SA is the Situational Awareness Rating Technique (SART) (Taylor, 1990). SART is a questionnaire method that concentrates on measuring the operator's knowledge in three areas: 1) demands on attentional resources, 2) supply of attentional resources, and 3) understanding of the situation (Figure 7 and Table 4). The reason SART measures three different components (there is also a 10-component version) is that the SART developers feel, like workload, SA is a complex construct; therefore, to measure SA in all its aspects requires separate measurement dimensions. Because information processing and decision making are inextricably bound with SA (since SA involves primarily cognitive rather than physical workload), SART has been tested in the context of Rasmussen's Model of skill-, rule-, and knowledge-based behavior. Selcon and Taylor (1989) conducted separate studies looking at the relationship between SART and rule- and knowledge-based decisions, respectively. The results showed that SART ratings appear to provide diagnosticity in that they were significantly related to performance measures of the two types of decision making. Early indications are that SART is tapping the essential qualities of SA, but further validation studies are required before this technique is commonly used.

		LOW						**HIGH**
		1	2	3	4	5	6	7
DEMAND	Instability of Situation							
	Variability of Situation							
	Complexity of Situation							
SUPPLY	Arousal							
	Spare Mental Capacity							
	Concentration							
	Division of Attention							
UNDERS	Information Quantity							
	Information Quality							
	Familiarity							

Figure 7. SART Scale

Table 4
Definitions of SART Rating Scales

Demand on Attentional Resources	
	Instability: Likelihood of situation changing suddenly.
	Complexity: Degree of complication of situation.
	Variability: Number of variables changing in situation.
Supply of Attentional Resources	
	Arousal: Degree of readiness for activity.
	Concentration: Degree to which thoughts bear on situation.
	Division: Amount of division of attention in situation.
	Spare Capacity: Amount of attention left to spare for new variables.
Understanding of the Situation	
	Information Quantity: Amount of information received and understood.
	Information Quality: Degree of goodness of information gained.

from Taylor and Selcon (1991, p. 10)

Strengths and limitations - SART is a subjective measure and, as such, suffers from the inherent reliability problems of all subjective measures. The strengths are that SART is easily administered and was developed in three logical phases: 1) scenario generation, 2) construct elicitation, and 3) construct structure validation (Taylor, 1989). SART has been prescribed for comparative system design evaluation (Taylor and Selcon, 1991). SART is sensitive to differences in performance of aircraft attitude recovery tasks and learning comprehension tasks (Selcon and Taylor, 1991; Taylor and Selcon, 1990). However, Taylor and Selcon (1991) state "There remains considerable scope for scales development, through description improvement, interval justification and the use of conjoint scaling techniques to condense multi-dimensional ratings into a single SA score" (p. 11). These authors further state that "The diagnostic utility of the Attentional Supply constructs has yet to be convincingly demonstrated" (p.12).

Data requirements - Data are on an ordinal scale; interval or ratio properties cannot be implied.

Thresholds - The data are on an ordinal scale and must be treated accordingly when statistical analysis is applied to the data. Non-parametric statistics may be the most appropriate analysis method.

SART was developed for British pilots. Americans have different connotations to some SART words.

LESSON NINE

Perceived imminence of death is a great SA enhancer. Firefighters who have been in rapidly deteriorating conditions experience heightened sensations: lights brighter, details clearer, sounds louder, smells more overpowering. With this often comes the perception that all action is in slow motion. One exhausted hero anchoring down a fire hose heard the crackling in an overhead pipe, smelled the first outpouring of gas, and saw the fireball forming and rolling towards three firefighters trying to free a man trapped under debris. The hero used the stream of water to deflect the fireball all in less than one second.

Figure 8. Hero's SA Saved Four Men

LESSON TEN

Building schema is critical to SA for humans and computers. Schema are patterns that help cut through data to the information. More experienced personnel typically have better schemas because they've been able to test their schema against more data. As Federico (1997) stated "Schema-driven decision making emphasizes the indispensability of (a) situational assessment in naturalistic settings and (b) knowledge base, past experience, event sequence, and similarity recognition as cognitive components of situation assessment" (pp. 149-150). The author concluded from the performance of naval officers reviewing tactical situations that "Participants' implicit beliefs in, automatic adherence to, and unconscious use of the schema-driven decision-making process are directly reflected in their performance of experimental tasks necessitating situation assessment" (p. 156).

In a recent ride on a transport aircraft, the loadmaster stopped a test engineer from measuring tensile strength of one of the load tiedowns. When asked why, the loadmaster replied, "Hear that hum? The aircraft is experiencing turbulence that has that chain humming at its resonant frequency." " So," asked the test engineer. "So that chain's going to snap." "Never," replied the test engineer. "You buying?" The test engineer lost $17 in drinks and a bit of his arrogance after the chain snapped and put a two inch dent in the cargo container against which the test engineer had been leaning during his measurements.

LESSON ELEVEN

SA, trust, and workload all must be considered during an evaluation. A young engineer happily reported that all pilots using the new sensor fusion system had destroyed all the threats. The engineer's boss was a bit skeptical. In his 25 years in the company he had never seen *every* pilot destroy *every* target. So the next day he decided to see for himself. Sure enough - *every* pilot, *every* threat. He was beginning to feel very pleased with the new system until he heard one of the pilots say, "Got every threat. Too bad I lost the airplane." The manager dashed over "Lost the airplane?" "Yep. But the young engineer told us to concentrate on destroying the threats, that nothing else would be scored. We couldn't handle destroying threats and saving our airplane so ..." Emphasizing SA may cause operators to have high workloads too high to perform other tasks well.

In a second system, battle management center operators were given a new decision aid designed to cull through vast amounts of data. The previous decision aid had missed critical information. Therefore, the operators didn't trust the new decision aid. They were told that they had to use the decision aid and they knew the importance of their jobs so they culled through the total data set as well as processed the information provided by the decision aid. Because they didn't trust the decision aid, they tried to do its work in parallel with the rest of their duties. Workload increased. Performance decreased and the decision aid was deemed a failure.

CONCLUSION

Eleven lessons in using SA to evaluate sensor fusion, decision aiding, and adaptive automation systems have been presented. These lessons fall into categories the importance of: 1) understanding the interaction of SA with workload, performance, and trust; 2) measuring at all three levels of SA; 3) considering the human; and 4) applying of good design engineering techniques.

REFERENCES

Adams, M.J. and Pew, R.W. Situational awareness in the commercial aircraft cockpit - A cognitive perspective. Proceedings of the Nineth IEEE/AIAA/NASA Digital Avionics Systems Conference. New York: Institute of Electrical and Electronics Engineers, Inc., 1990, p. 519-524.

Carretta, T.R., Perry, D.C., and Ree, M.J. Prediction of situational awareness in F-15 pilots. International Journal of Aviation Psychology. 6(1), 21-41; 1996.

Derrick, W.A. (1985). The dissociation between subjective and performance-based measures of operator workload. National Aerospace and Electronics Conference, 1020-1025.

Endsley, M.R. A construct and its measurement: the functioning and evaluation of pilot situation awareness (NORDOC 88-30). Hawthorne, CA: Northrop Corporation; 1988c.

Endsley, M.R. A methodology for the objective measurement of pilot situation awareness. Presented at the AGARD Symposium on Situation Awareness in Aerospace Operations. Copenhagen, Denmark; October 1989.

Endsley, M.R. Design and evaluation for situation awareness enhancement. Proceedings of the 32nd Annual Meeting of the Human Factors Society. 97-101; 1988b.

Endsley, M.R. Predictive utility of an objective measure of situation awareness. Proceedings of the Human Factors Society 34th Annual Meeting (pp. 41-45), Santa Monica, CA: Human Factors Society, 1990b.

Endsley, M.R. Situation awareness and workload: flip sides of the same coin. Proceedings of the Seventh International Symposium on Aviation Psychology ; Volume 2, Columbus, OH, Ohio State University, 1993, p. 906-911.

Endsley, M.R. Situation awareness in dynamic human decision making: theory and measurement (NORDOC 90-49). Hawthorne, CA: Northrop Corporation, 1990a.

Endsley, M.R. Situation Awareness in dynamic systems. In R.M. Taylor (Ed.) Situational awareness in dynamic systems (IAM Report 708). Farnborough, UK: Royal Air Force Institute of Aviation Medicine, 1991b.

Endsley, M.R. Situational awareness global assessment technique (SAGAT). Proceedings of the National Aerospace and Electronics Conference. New York, Institute of Electrical and Electronic Engineers, Volume 3, p. 789-795; 1988a.

Federico, P. An empirical examination of metacognitive models of situation assessment. Human Factors, 1997, 39(1), 149-157.

Fracker, M.L. A theory of situation assessment: implications for measuring situation awareness. Proceedings of the 32nd Annual Meeting of the Human Factors Society (pp. 102-106). Anaheim, CA; 1988.

Fracker, M.L. Attentional allocation in situation awareness. Proceedings of the 33rd Annual Meeting of the Human Factors Society (pp. 1396-1399). Denver, CO; 1989.

Fracker, M.L. and Davis, S.A. Measuring operator situation awareness and mental workload. Proceedings of the Fifth Mid-Central Ergonomics/Human Factors Conference. Dayton, OH; 23-25 May 1990.

Fracker, M.L. and Vidulich, M.A. Measurement of situation awareness: a brief review. In R.M. Taylor (Ed.) Situational awareness in dynamic systems (IAM Report 708). Farnborough, UK: Royal Air Force Institute of Aviation Medicine, 1991.

Mosier, K.L. and Chidester, T.R. Situation assessment and situation awareness in a team setting. In R.M. Taylor (Ed.) Situational awareness in dynamic systems (IAM Report 708). Farnborough, UK: Royal Air Force Institute of Aviation Medicine, 1991.

Sauter, N.B. and Woods, D.D. Situational awareness: a critical but ill-defined phenomenon. International Journal of Aviation Psychology. 1(1), 45-57; 1991.

Selcon, S.J. and Taylor, R.M Decision support and situational awareness. In R.M. Taylor (Ed.) situational awareness in dynamic systems (IAM Report 708). Farnborough, UK: Royal Air Force Institute of Aviation Medicine, 1991.

Selcon, S.J. and Taylor, R.M. Evaluation of the situational awareness rating technique (SART) as a tool for aircrew systems design. AGARD Conference Proceedings No. 478. Situational Awareness in Aerospace Operations, p. 8, Neuilly-sur-Seine, France; 1989.

Selcon, S.J., Taylor, R.M, and Koritsas, E. Workload or situational awareness?: TLX vs SART for aerospace systems design evaluation. Proceedings of the Human Factors Society 35th Annual Meeting (pp. 62-66) San Francisco, CA; 1991.

Taylor, R.M. Situational awareness: aircrew constructs for subject estimation (IAM-R-670); 1990.

Taylor, R.M. Situational awareness rating technique (SART): the development of a tool for aircrew systems design. Proceedings of the NATO Advisory Group for Aerospace Research and Development (AGARD) Situational Awareness in Aerospace Operations Symposium (AGARD-CP-478); October 1989, p. 17.

Taylor, R.M. and Selcon, S.J. Subjective measurement of situational awareness. In R.M. Taylor (Ed.) Situational awareness in dynamic systems (IAM Report 708). Farnborough, UK: Royal Air Force Institute of Aviation Medicine, 1991.

Taylor, R.M. and Selcon, S.J. Understanding situational awareness. Proceedings of the Ergonomics Society's 1990 Annual Conference. Leeds, England; p. 3-6, 1990.

Wickens, C.D. and Liu, Y. (1988). Codes and modalities in multiple resources: A success and a qualification. Human Factors, 30, 599-616.

Yeh, Y. and Wickens, C.D. (1988). Dissociation of performance and subjective measures of workload. Human Factors, 30, 111-120.

NAVY TACAIR RADAR WARNING RECEIVER - CAPABILITIES AND LIMITATIONS AS A SITUATION AWARENESS SENSOR

Dr. Douglas Sticht	Mr. Peter Mulham	Mr. Stuart Johnson
Naval Air Warfare Center	Naval Air Warfare Center	Naval Air Warfare Center
Weapons Division	Weapons Division	Weapons Division
TACAIR EW IPT Leader	RWR Systems Branch	F/A - 18 WSSA
Pt Mugu, CA	Pt Mugu, CA	China Lake, CA

This paper discusses the type and quality of data that today's deployed, in-production, and developmental tactical aircraft (TACAIR) radar warning receivers (RWRs) bring to aircrew situation awareness (SA). While similarities and contrasts are made, it is not a comparison of RWRs. In the same context, it is not a discussion of electronic warfare suites, although the RWR can provide SA to jammers which have no capability to determine whether their received threat is forward or aft of the aircraft, and it also provides SA for expendables dispensing which can be optimized using bearing to the threat. The U.S. Navy's fleet-deployed TACAIR RWR is the AN/ALR-67(V)2, and the fleet's helicopter RWR is the AN/APR-39(V)1. These systems provide real-time, on-the-scene, on-board information for SA to air crew. However, their ultimate performance is limited by ambiguities in frequency and pulse repetition interval (PRI) among threat emitters, and it is bounded by achievable direction of arrival accuracy because of system cost and installation constraints. New platforms and mission systems upgrades combine the RWR information with other on-board sensor data, mission planning data, and tactical data link information to present a fused SA composite to the air crew. With properly integrated sensor data, controls, audio, and displays, the air crew has the most precise, accurate, and timely SA available for real-time tactical decisions. The RWR's ability to define the threat directly impacts those decisions, and can turn potentially damaging situations around to tactical victory.

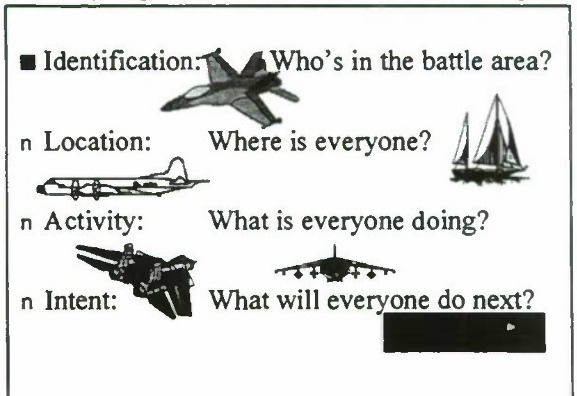


Figure 1. Situation Awareness In This Paper.

For purposes of this paper, SA is defined as "knowing what is going on," and will specifically exclude discussion of potential actions taken as a result of that knowledge. We shall examine the TACAIR RWR's contribution to "knowing what is going on" in four informational categories as depicted in Figure 1:

• Identification: who is in the battle area? This may include wingman, coalition force members, hostiles, and non-combatants.

• Location: where is everyone? There may be a mission need to target and deliver ordnance, to escort other aircraft, or to deny hostile access to airspace.

• Activity: what is everyone doing? This includes parameters such as range, speed, and heading as well as whether engaging a weapons system as a hostile might do, or simply fishing as a civilian boat might be doing.

• Intent: what is everyone going to do next? This is perhaps the most challenging information to obtain, yet it can be most decisive in battle.

The RWR brings crucial, real-time threat information to the situation awareness picture, and in general the more modern the RWR, the better the information in several contexts. RWR information provided is summarized in Figure 2. For identification, the RWR is limited to radio-frequency (RF) emitters which are either explicitly in the RWR's knowledge base (known as mission data file, user data file, or some similar term), or emitters whose waveforms behave similarly to weapons system waveforms, in which case a symbol indicating an unknown emitter will be passed along to the air crew as a potential threat.

To locate emitters and display direction of arrival (DOA) to the air crew, Navy deployed TACAIR RWRs are implemented as amplitude comparison systems with generally four antennas boresighted along

four quadrants in the wing plane. Thus, they provide signal DOA (bearing) of emitters in the azimuthal (wing) plane only. Although RWRs provide a substantial amount of elevation detection capability, no direction of arrival may be calculated or displayed in elevation. RWRs have no way of measuring or displaying range, and speed and heading is best deduced by the operator from the on-board radar.

Navy RWRs provide SA activity information if the emitter parameters match their respective mission data library (knowledge base), or, if unknown, are generically capable of supporting weapons fire control. More precise knowledge of the emitter's activity is available if the RWR is able to recognize and correlate multiple signals from the same threat weapons system.

A good RWR is an excellent intent sensor. As the fire control system's signals change, a good RWR is able to deduce that a weapon system has progressed from acquisition mode to target tracking mode to modes which support missile attack. The ideal is to be able to look for "smoke in the air" before it is there.

Navy fleet helicopters today utilize the AN/APR-39(V)1 RWR. This warning system provides pulse threat detected audio directly to the pilot who then must identify the threat by its tone characteristics and estimate its approximate range through signal intensity and clarity. There is an azimuthal display on which a direction of arrival strobe appears in order to indicate approximate bearing to the threat fire control signal, while signals in the lowest radio frequency band are simply indicated by a light. In areas of relatively few threat signals, the astute pilot with enough time to focus on the signals may deduce threat activity and intent by correlating two or more signals from the threat weapons system. If continuous-wave (CW) signal detection and warning capability is required, the AN/APR-44(V) is added to the aircraft, such as in the UH-1N and the AH-1W.

What An RWR Brings to SA

- Identification: RF Emitters
- Location: Azimuth In Wing Plane
- Activity: Fire Control Radar Systems
- Intent: Weapon Control Waveforms

-> **Real-Time, On-Board, On-The-Scene**

Figure 2. RWR Information Provided.

Recognizing a need to detect a greater number of threats across broader frequency coverage, to provide for reprogrammable mission data libraries, and to interface with a wide range of associated avionics systems, the AN/APR-39A(V)2 was developed as a replacement for the AN/APR-39(V)1 as well as for MV-22 and other slower fixed-wing aircraft (e.g., KC-130), and a Navy production contract was awarded in 1996. This system is markedly advanced over the AN/APR-39(V)1 in that it features much greater total RF bandwidth, built-in CW capability, high pulse rate tracking and processing for pulse doppler (PD) capability, digital signal processing for threat identification, and alphanumeric graphical display with azimuth DOA indication. Certain waveforms associated with multiple target track weapons systems and designated as low probability of intercept (LPOI) waveforms are

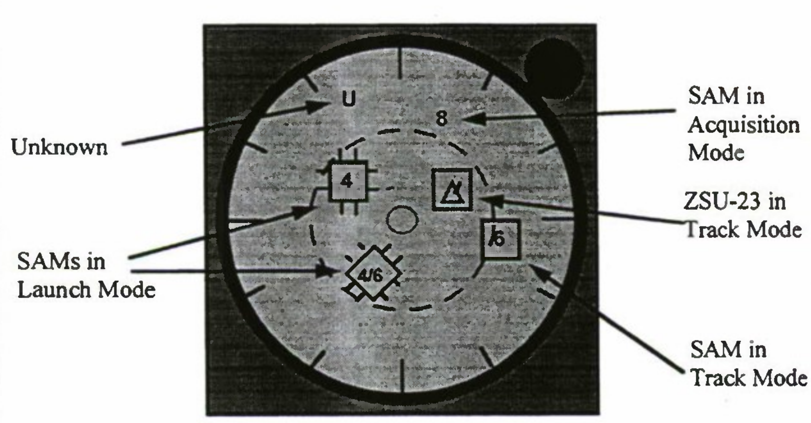


Figure 3. AN/APR-39A(V)2 Display Example.

programmed into the receiver scanning schedule, digital signal processor, and emitter threat library. There is a narrowband tunable receiver channel available to identify and track CW and PD threats, and this channel is also tasked by the signal processor when more precise frequency measurement is necessary to resolve ambiguity between two emitters. The AN/APR-39A(V)2 uses amplitude comparison to calculate DOA, similarly to the AN/APR-39(V)1, and the alphanumeric graphical symbology for a particular emitter is displayed on the azimuthal indicator at the received bearing, relative to aircraft wings level. Emitters are identified and displayed symbolically, with voice warnings provided to indicate that a new emitter has been displayed, an air interceptor has been displayed, or a weapons system has changed operational mode, for instance by change of emitter waveform and/or additional emitter signal from the same weapons system. Weapons system activity and intent is monitored through changes and/or additions to emitter waveforms from a weapons system being tracked, and a set of graphical alphanumeric symbols and voice warnings is used to display activity and intent. The AN/APR-39A(V)2 can display not only its own RWR data, but also AN/AAR-47 missile warning system information and AN/AVR-2 laser warning system data, and audio tones for each system are provided. The AN/APR-39A(V)2 display is shown in the example in Figure 3, and all the warning information may be transmitted via a MIL-STD-1553B data bus.

Today's deployed fixed-wing fighter and attack RWR is the AN/ALR-67(V)2. The AN/ALR-67(V)2 is quite similar in function to the AN/APR-39A(V)2, however, its mission requirements dictate a greater system pulse throughput because of its denser signal operating environment, and the AN/ALR-67(V)2 can accommodate more than twice as many emitters in its mission data file. Figure 4 summarizes SA data available from the AN/ALR-67(V)2 and the AN/APR-39A(V)2 RWR systems. The AN/ALR-67(V)2 does not provide voice warnings; a sequence of audio tones is generated, depending on whether a new emitter has been displayed, an airborne interceptor emitter has been displayed, or an emitter has changed. While the AN/ALR-67(V)2 has an associated display as depicted in the example in Figure 5, the preference in modern fighter and attack aircraft is to route the display information to primary aircraft displays such as the multi-purpose display and heads-up display, where the RWR data are combined with other data useful and necessary to the pilot. Threat display is available on multi-purpose and heads-up displays on F/A-18, F-14, and AV-8B, and examples are discussed later in this paper.

AN/APR-39A(V)2 & AN/ALR-67(V)2

■ Identification:	Pulse, CW, PD, LPOI Emitters
	Alphanumeric Graphical Symbols
n Location:	Amplitude Comparison DOA
	Azimuthal Indicator Display
n Activity:	Displayed Symbology
	Voice Warnings (APR-39A(V)2)
	Audio Tone Set (ALR-67(V)2)
n Intent:	Symbology/Symbology Changes
	Voice Warnings (APR-39A(V)2)
	Tone Changes (ALR-67(V)2)

Figure 4. SA Data From Modern RWRs.

Modern so-called "fourth generation" airborne interceptor (AI) pulse-doppler radars utilize very similar pulse repetition rates (PRIs) and nearly identical frequencies. As a result, there are instances when the usual RWR discriminators of pulse repetition interval within a relatively wide frequency band cannot unambiguously resolve and display received signals. One approach to resolving these ambiguities is being implemented by more precisely defining PRI ranges ("sliver" PRIs) for emitters in the newer RWR data libraries. Modern AI threat radars also generate tremendously complex sequences of waveforms as they perform their multiple functions via multiple operating modes. Tracking one target while scanning for others, and measuring range to one target while searching for others are two examples of such radars' complex functionality. Implementing those functions can involve changing PRI, interleaving sequences of different PRIs, and switches in frequency. To detect and track the associated waveforms, the RWR needs a fast and precision frequency measurement, a flexible frequency scanning schedule, a very capable signal processor, and precision emitter data. If the radar emitter being tracked switches from a mode which was formerly unique to a highly ambiguous mode, it is imperative that the RWR sustain the identification.

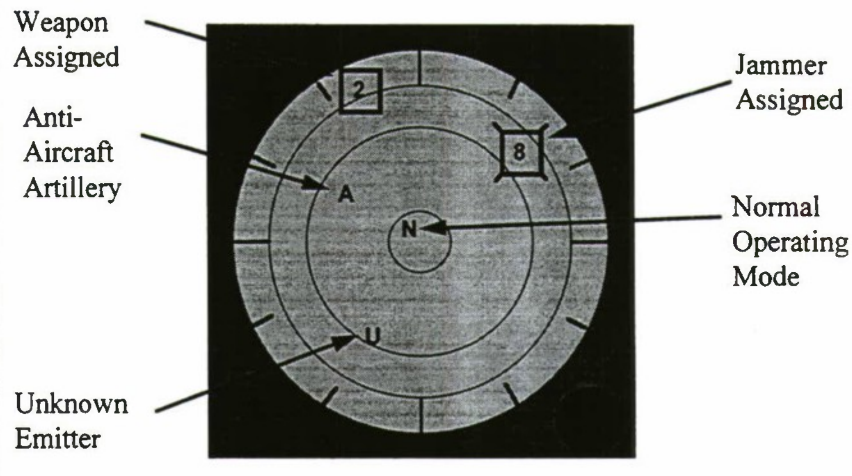


Figure 5. AN/ALR-67(V)2 Display Example.

Responding to threat AI advancements incorporating pulse doppler radar and fire control, the Navy initiated the AN/ALR-67(V)3 program. This RWR provides capability for longer detection ranges,

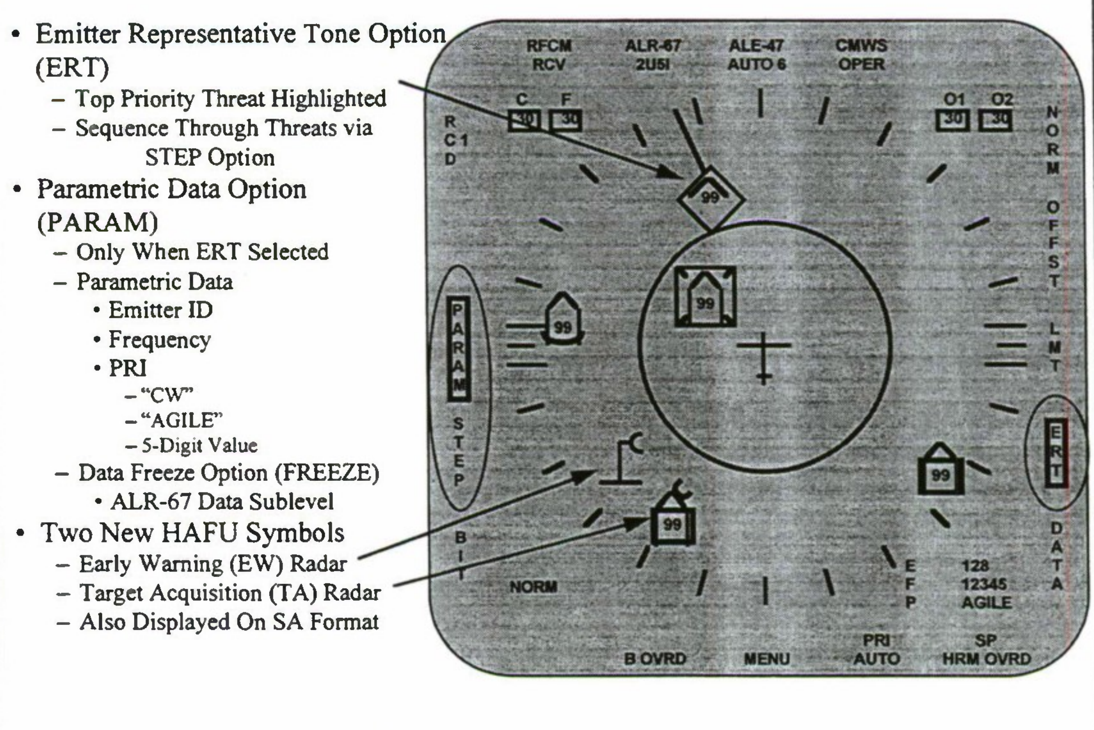


Figure 6. F/A-18 AN/ALR-67(V)3 Symbology.

more ability to resolve ambiguities, and capability to operate in a higher pulse density signal environment. It achieves this performance through higher system pulse detection sensitivity, precision frequency and pulse width measurement on each pulse, and precision received amplitude measurement. Higher sensitivity provides an additional benefit of being able to achieve better DOA at weak signal levels because all four receiver channels may be employed in the DOA calculation whereas for a less sensitive system only a single quadrant receiver may be able to detect the signal. The AN/ALR-67(V)3 is completing development, and performance comparison with the existing AN/ALR-67(V)2 TACAIR RWR will be part of the operational evaluation.

Figure 6 depicts AN/ALR-67(V)3 symbology in development for the F/A-18 multi-purpose display (MPD), a full color display in the aircraft. The top priority displayed threat, determined from the mission data file, is highlighted, while various symbols depict hostile, ambiguous, friendly, and unknown (HAFU) emitters. There are control options to present emitter-representative tones instead of the earlier AN/ALR-67(V)2 tone set. Jammer and weapon assignment enhancements are provided as with the AN/ALR-67(V)2. The pilot has the option to display more emitter parametric data than ever before available, and new symbols are provided for threat types tracked by the more capable AN/ALR-67(V)3 RWR.

For the F/A-18 there is in development a combined-sensor situation awareness (SA) display integrated with system controls and audio tones. The first version of this display is in final evaluation for

- Plan View
  - A/A MSI Trackfiles
  - ALR-67 MSI Trackfiles
  - Attack Region Cues (± 70°)
  - Sensor Field-of-View Cues
  - A/A Waypoint
  - 1/2 Range Ring
- Sensor Sublevel Option (SENSR)
  - A/A MSI Contributor Selection
- Digital Map Option (MAP)
  - Displays Digital Map Underlay
  - Map For 40, 20, 10, 5 NM Scales
- Range Scale Option (SCL/###)
  - 320, 160, 80, 40, 20, 10, 5 NM
- Decenter Mode Option (DCNTR)
  - Selects Decentered Format
- Trackfile Step Box Option (STEP)
  - Boxes Link-4 Trackfiles In Sequence

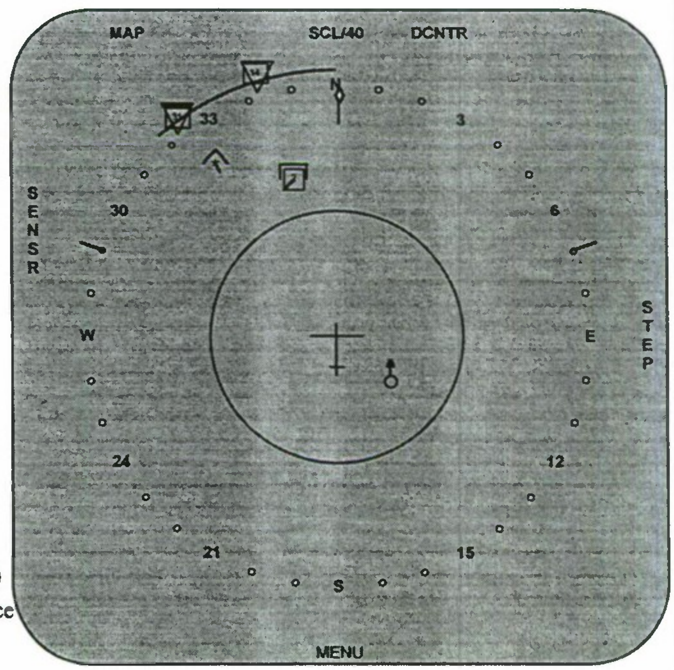


Figure 7. F/A-18 MPD SA Format for F/A-18 Block 11C.

the C and D models with mission software set 11C. The fundamentals of this display are shown in Figure 7, and there are many options and features available to include underlay display of the digital map for a selection of range scales, and a decentered format in case more detail in the forward attack region is desired.

As part of the Integrated Defensive Electronic Countermeasures (IDECM) program, the F/A-18E/F is refining the man-machine interface to incorporate missile warning sensor directional strobes, countermeasures subsystems status, and controls for countermeasures response modes, including silent. Navigation options include marked way points and waypoint steering cues with sequence numbers.

Multi-sensor/multi-source integration is provided wherein correlated track files are presented as single tracks, and source information is available at a glance for each displayed track. An example of the F/A-18E/F SA format is shown in Figure 8. The SA format provides up to 10 trackfiles which have been correlated through multiple sensor data fusion. Colors are used to differentiate friendly, nonlethal, lethal,

- Up To 10 Trackfiles Displayed
  - Dugout for No-Angle-Tracks
  - Priority Based On RWR Lethality and Degree Of Ring Penetration
  - Angle-Only Tracks (AOT) Displayed Outside Compass Rose
  - Displayed as AOT Until Range Confidence < 10 NM
- Threat ID Characters In Center
- Missile Warning Sensor Strobe
- Static Enhancement Symbol
  - SAM, AAA, and/or Ship-Based
  - Removed via DCLTR
- Dynamic Enhancement Symbol
  - HARM (Box)
  - ECM In Progress (Whiskers)
- Color Indicates Lethality/ Allegiance
  - Green = Friendly
  - Yellow = Nonlethal
  - Red = Lethal
  - Flashing Red = Critical/MA/ML

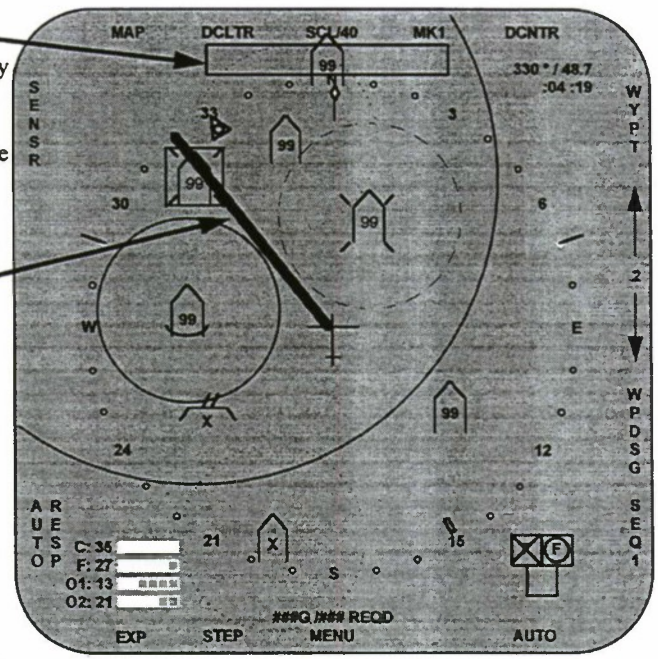


Figure 8. F/A-18E/F Multi-Purpose Display Expanded Multi-Source Trackfiles

and critical status, and for the top five priority threats, circular rings are provided which indicate whether the threat was obtained from prebriefed data, RWR only, or from correlated multiple sensors. Finally, the threat ring sizes are related to threat lethal range. Even with multi-source integration and fused track displays, there is need for the decluttering options in order to focus on specific aspects of the mission at any particular time. Avoiding pop-up threats and providing weapon correct targeting information can each demand focus and the convenience of a decluttered display, although critical threat information, such as missile warning sensor directional strobe accompanied by voice warning, jammer failure, and other cautions and advisories have high precedence for display as well as audio.

The heads-up display, shown in Figure 9, includes missile warning sensor directional strobe, a limited number of threats, and countermeasures response cueing which indicates automatic mode response in process, semi-automatic mode assent request, or cue for single response flare, chaff, towed decoy enable, expendable jammer, etc. Under sponsorship of the IDECM program, these F/A-18E/F controls and displays are prototyped in a simulated cockpit at McDonnell-Douglas Aerospace, and have been formally evaluated four times by a team consisting of representatives of Navy developers, F/A-18E/F integrated test team, VX-9, Top Gun, Strike Warfare Center, and Pacific fleet fighter wing. The controls and displays are improved after each formal evaluation. The most recent evaluation, known as SimEval #4, was conducted on 12 March 1997, and included 14 participating aircrew "flying" the configuration known as 16E. Each participant filled out a 24-question evaluation which addressed SA format, countermeasures response management, EW format, warnings and advisories, and the simulation scenario. All were quite well received, and the few remaining suggestions for improvement will be considered for incorporation by the system design advisory group.

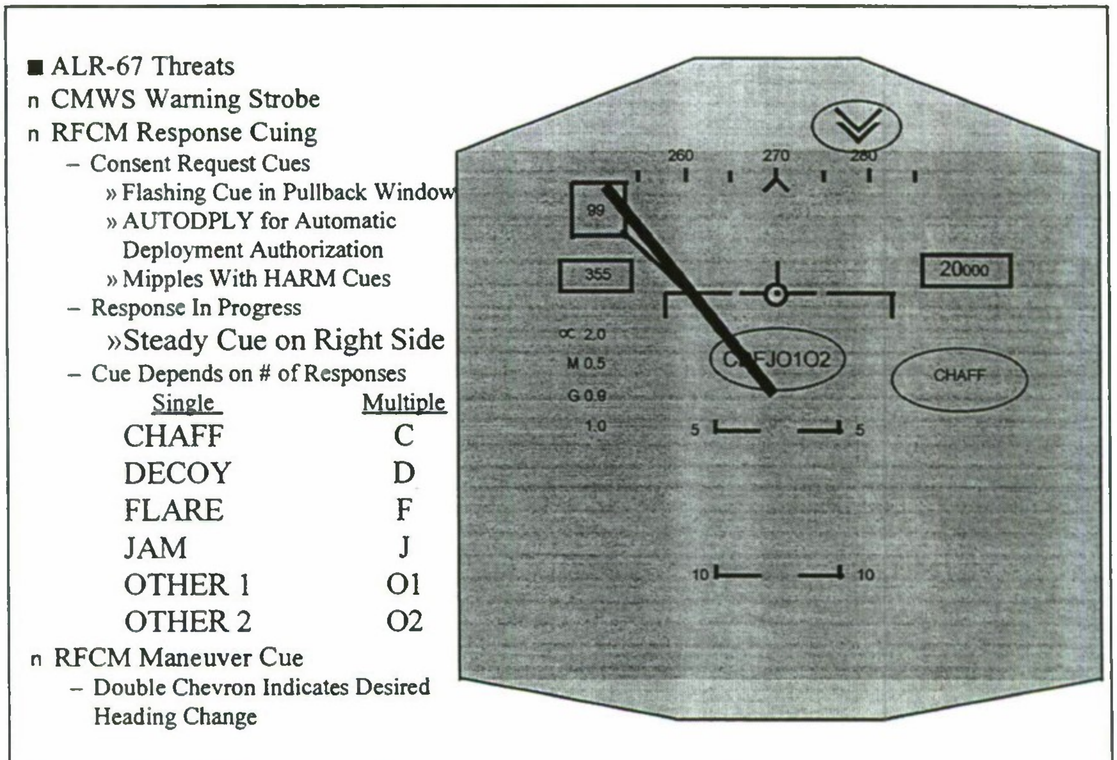


Figure 9. F/A-18E/F Heads-Up Display Illustrating SA and MSI Features.

In conclusion, it is seen that the modern RWR systems provide real-time, on-the-scene, on-board information for threat situational awareness (SA) to air crew. They can also provide useful information to optimize countermeasures. New platforms and mission systems upgrades are combining the RWR information with other on-board sensor data, mission planning data, and tactical data link information to present a fused SA composite of facts to the air crew. With properly integrated sensor data, controls, audio, and displays, the air crew has the most precise, accurate, and timely SA available for real-time tactical decisions. The RWR's ability to define the threat directly impacts those decisions, and can turn potentially damaging situations around to tactical victory.

This page intentionally blank.

REAL-TIME FUSION OF LOW-LIGHT VISIBLE AND FLIR IMAGERY: COLOR NIGHT VISION FOR ENHANCED SITUATIONAL AWARENESS

A.M. Waxman, M. Aguilar, D.A. Fay,
J.P. Racamato, J. Carrick, M. Seibert, E.D. Savoye
MIT Lincoln Laboratory
Lexington, MA 02173

Both intensified visible and thermal infrared (IR) imagery are commonly utilized to enhance situational awareness during night operations on the ground and in the air. It is well known that these imaging modalities provide complementary information to the user, and pilots commonly try to mentally fuse this information. It is now accepted that a real-time image fusion system, capable of exploiting the wide dynamic range imagery provided by night sensors, would be of great value in the cockpit and on the ground. Research at MIT Lincoln Laboratory since 1993 has been directed at this goal of real-time visible/IR fusion with color presentation, as well as entirely solid state color fusion night vision devices. This paper reports substantial progress towards these goals.

We describe an apparatus and methodology to support real-time color imaging for night operations. Registered imagery obtained in the visible through near-IR band is combined with thermal IR imagery using principles of biological opponent-color vision. Visible imagery is obtained using a Gen III image intensifier tube fiber-optically coupled to a conventional Charge Coupled Device (CCD), and thermal IR imagery is obtained using an uncooled thermal imaging array, the two fields of view being matched and imaged through a dichroic beam splitter. Realistic color renderings of a variety of night scenes are illustrated. We also demonstrate both grayscale and color fusion of intensified-CCD/FLIR imagery recorded during night helicopter flight, obtained from the U.S. Army Night Vision and Electronic Sensors Directorate (NVESD). Computations are carried out in real-time on commercially available C80 Digital Signal Processor (DSP) boards.

Our color image fusion has also been applied to Canadian imagery to demonstrate smokescreen penetration, and to Dutch surveillance imagery taken under very low visible/thermal contrast conditions. Human perception experiments have now been conducted at three independent laboratories, all of which indicate improved situational awareness by subjects viewing our color fused imagery (as compared to original visible or IR imagery, or alternative fusion products).

Progress in the development of a low-light sensitive CCD imager with high resolution and wide intrascene dynamic range, operating at 30 frames/sec is also described. Example CCD imagery obtained under controlled illumination conditions (without an intensifier tube), from full moon down to overcast starlight conditions, processed by our adaptive dynamic range algorithm is shown.

The combination of low-light visible CCD imager and uncooled thermal IR microbolometer array in a single dual-band imager, with portable image processing computer implementing our neural net algorithms, and color LCD display, yields a compact integrated version of our system in the form of a solid state color night vision device for use on the ground, sea, and possibly for helicopter operations as well. The systems described here have application to a large variety of military operations and civilian needs.

(Reprint of executive summary; formal paper not available)

This page intentionally blank.

TOWARDS A JOINT HUMAN-MACHINE MODEL FOR SITUATION AND THREAT ASSESSMENT

Stéphane Paradis
Defence Research Establishment
Valcartier, Val-Bélair, Qué.,
CANADA

Bruce A. Chalmers
Defence Research Establishment
Valcartier, Val-Bélair, Qué.,
CANADA

Pierre Bergeron
Lockheed Martin Canada
Montréal, Qué.,
CANADA

ABSTRACT

At the heart of the combat system, is a command and control system (CCS) located in the operational room of the ship, by which commanders can plan, direct, control, and monitor any operation, for which they are responsible, to defend their ship and fulfill their mission. Technological advances in threat technology, the increasing tempo, and diversity of open-ocean and littoral scenarios, and the volume and imperfect nature of data to be processed under time-critical conditions pose significant challenges for future shipboard CCS. This emphasizes the need for naval warships to be fitted with an efficient combat system featuring a real-time, joint human-machine oriented decision support capability integrated into the ship's command and control system. This decision support system consists of a multi-source data fusion (MSDF) capability, a situation and threat assessment (STA) capability, and resource management (RM) capability (for weapons and sensors). These capabilities are required to support combat operators in their command and control activities, to ensure own ship survival, and to increase the probability of mission success.

In order to address these challenges, the Data Fusion & Resource Management Group in the Decision Support Technologies Section at Defence Research Establishment Valcartier (DREV) has undertaken the investigations of design concepts for real-time decision support system (DSS), called MSDF/STA/RM system for the Canadian Patrol Frigate (CPF) in order to improve its performance against current and future threats. The research and development work is conducted following a human-centered philosophy and taking into consideration situation awareness concepts.

This paper presents the preliminary research work on a joint human-machine model for STA. The aim of the research is to explore the problem of fusing information provided by MSDF with that from external environmental sources of information in order to determine the probable situation explaining the presence, status, allegiance, and intentions of the observed entities and to prioritize the identified threats. Although this is not an easy thing to do, the real challenge remains in the implementation of a STA system that will complement or support the decision-makers. Such a system must be designed so that the operator will trust and use it. The ultimate goal should be to implement a system that executes the STA process faster and more efficiently by a human-machine combination working in synergy.

A first cut of a model for STA is presented. The proposed model evolved from a three level descriptive model of situation awareness, where the first level is related to perception of the elements in the current situation, the second level is about comprehension of the current situation, and the last level deals with projection of the future situation states. This model yields a high-level functional decomposition of a multilevel STA process. Finally, a description of our STA-implementation-under-construction is presented.

INTRODUCTION

At the heart of the combat system is a command and control system (CCS) located in the operational room of the ship, by which commanders can plan, direct, control, and monitor any operation, for which they are responsible, to defend their ship and fulfill their mission. Technological advances in threat technology, the increasing tempo, and diversity of open-ocean and littoral scenarios, and the volume and imperfect nature of data to be processed under time-critical conditions pose significant challenges for future shipboard CCS. The Data Fusion and Resource Management Group at Defence Research Establishment Valcartier (DREV) and the research and development department at Lockheed Martin Canada (LMC) have, for several years now, been investigating concepts to continuously fuse data from a ship's sensors and other sources, dynamically maintain a tactical situation picture, and support response to actual or anticipated threats. The two organisations are now jointly involved in a project to design,

develop, and implement a real-time decision support system (DSS) that can be integrated into the ship's CCS to support operators in their tactical decision making and action execution activities, with particular focus on application to the mid-life upgrade of the Canadian Patrol Frigate (CPF) expected early in the next century.

This paper is concerned with situation and threat assessment (STA) support capability of the DSS for above water warfare (AWW) context. The aim of the paper is to review background information and fundamentals on STA, define new concepts, and present a first cut of a generic model influenced by the human's mental processing. This model for STA is built from a three-level descriptive model of situation awareness, in which the first level is concerned with perception of the elements in the environment, the second level concerns comprehension of the current situation, and the last level deals with projection of future states of the situation. This model leads us to a high-level functional decomposition of a multi-level STA process. The short term goal is to support the human to assess the situation and the threat by the automation of some higher level cognitive processing currently performed by the human. The ultimate goal is aimed at the design and implementation of a joint human-machine system working in synergy that will complement and support the decision-makers and executes the STA process faster and more efficiently.

Finally, ongoing work in a collaborative DREV/LMC project which is implementing a real-time Blackboard and knowledge sources (intelligent agents) based on this model is briefly discussed.

This paper is organized as follows. Section 2 provides brief background information on the tactical command and control (C2) process. In section 3, technological and cognitive issues with respect to decision support are discussed. Fundamentals and requirements for STA are presented in sections 4 and 5 respectively. A first cut of the generic model is presented and described in section 6. Section 7 briefly describes the DREV/LMC project which is currently implementing a real-time Blackboard with knowledge sources (intelligent agents) based on the generic model. Section 8 provides conclusions.

COMMAND AND CONTROL PROCESS

Command and control (C2) is the process by which commanders can plan, direct, control, and monitor operation for which they are responsible. In a naval context, most tactical decision done within the ship's Operations Room, is made after completing a number of perceptual, procedural, and cognitive activities linked to the C2 process. The C2 process is a suite of periodic activities which mainly involves perception of the domain (environment), assessment of the tactical situation, decision making about a course of action and implementation of the chosen plan. Characteristics of this suite of activities were captured through Boyd's Observation-Orient-Decide-Act (OODA) loop illustrated in Fig. 1.

The C2 activities are performed by either the human, the machine, or a combination of both. A team

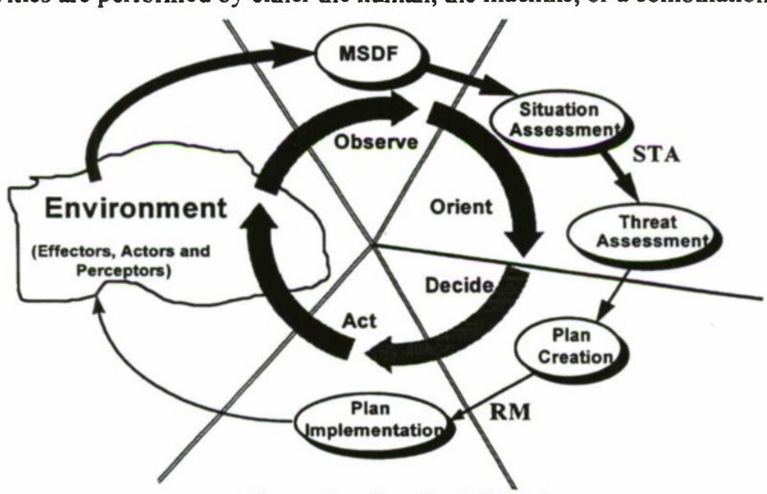


Figure 1. - Boyd's OODA Loop

of combat system operators interact with a command and control system (CCS) through consoles, aided by a number of other systems, to perceive the domain and build the maritime tactical picture (MTP). Then, the team will, based on the MTP and other sources of information (a priori knowledge, data link, intelligence reports, human experience, etc.), interpret and assess the tactical situation with the help of

decision support tools. The resulting assessment will increase the commander's situational awareness and put him/her in a better position to decide, in light of the mission's goals and the tactical situation, on a course of action to implement.

Military community says that the dominant requirement to counter the threat and ensure the survivability of the ship is the ability to perform C2 activities (OODA loop) quicker and better than their adversary. Therefore, the speed of execution of the OODA loop and the degree of efficiency of its execution are the keys of success of shipboard tactical operations. The reader can refer to [1-2] for a more detailed description of the C2 process.

DECISION SUPPORT SYSTEM

Technological advances in threat technology and the volume, rate, imperfect nature, and complexity of the information to be processed under time-critical conditions may exceed the ability of human operators to cope with the situation. Therefore, to address these challenges, operators need to be aided, while performing command and control (C2) activities, by an automated, real-time decision support system (DSS), called multi-source/situation and threat assessment/resource management (MSDF/STA/RM) system, that continuously fuses data from the ship's sensors and other sources, helps operators maintain a tactical situation picture, and supports their response to actual or anticipated threats. Therefore, the main role of a real-time DSS is to support operators in their tactical decision making and action execution activities. Situation and threat assessment (STA) support capability represents one dimension of the decision support technology which is addressed in the next section.

Providing combat system operators in the operations room with automated, real-time decision support raises a host of technological and cognitive issues that require careful consideration. Technological issues address the hardware and software aspects of automating the information processing, decision analysis and control capabilities of the embedded support system. Cognitive issues are concerned with the specifics of the various cognitive-level behaviors (e.g., perception, monitoring, planning, problem solving, and decision making) which the decision support system must exhibit and/or support for its users and the modes of human-computer interaction (HCI) between these users and the automated system via a HCI, including the content, structure, and form of such interactions. The reader can refer to [1-2] for a more complete and detailed discussion on technological and cognitive design issues related to DSS.

FUNDAMENTALS ON SITUATION AND THREAT ASSESSMENT

Situation and threat assessment (STA) corresponds to levels 2 and 3 of the Joint Director of Laboratories (JDL) Data Fusion (DF) model [3-4].

Data Fusion

DF is fundamentally, according to the JDL model, a process designed to manage (i.e., organize, combine, and interpret) data and information, obtained from a variety of sources, that may be required at any time by operators and commanders for decision making. It's an adaptive information process that continuously transforms the available data and information into richer information, through continuous refinement of hypotheses or inferences about real-world events, to achieve refined (and potentially optimal) kinematics and identity estimates of individual objects, and complete and timely assessments of current and potential future situations and threats (i.e., contextual reasoning), and their significance in the context of operational settings. The DF process is also characterized by the evaluation of the need for additional data and information sources, or modification of the process itself, to achieve improved results.

The process of DF may be viewed as a multilevel hierarchical inference process whose ultimate goal is to assess a mission situation and identify, localize, and analyze threats. However, not every DF application is responsible for all of these outputs. Some applications are only concerned with the position and identification of objects. Other applications are primarily oriented towards the situation and how it is evolving. Still, others focus on the threat and its possible impact upon achieving mission objectives. In addition, the DF function can be responsible for identifying what information is most needed to enhance its products and what sources are most likely to deliver this information.

Given these considerations, a complete DF system can typically be decomposed into four levels [3]:

- Level 1 - Multi-Source Data Fusion (MSDF);
- Level 2 - Situation Assessment (SA);
- Level 3 - Threat Assessment (TA); and,

• Level 4 - Process Refinement Through Resource Management (RM).

Each succeeding level of DF processing deals with a higher level of abstraction. Level 1 DF uses mostly numerical, statistical analysis methods, while levels 2, 3, and 4 of DF use mostly symbolic or Artificial Intelligence (AI) methods. Note that resource management in the context of level 4 fusion is mainly concerned with the refinement of information gathering process (i.e., sensor management). However, the overall domain of resource management also encompasses the management of weapon systems and other resources.

Situation Assessment

SA is a recent topic of research and is consequently an immature field in comparison with MSDF. This explains the existence of so many formal definitions of the concept of SA in the literature and highlights the imprecise and disparate interpretations in different papers. The DF sub-panel[2] of the JDL/TPC3 defines SA, also known as situation refinement, as the level two processing which develops a description or interpretation of current relationships among objects and events in the context of the operational environment. The result of this processing is a determination or refinement of the battle/operational situations. Key functions include :

1. Object Aggregation - Establishment of relationships among objects including temporal relationships, geometrical proximity, communications links, and functional dependence.
2. Event/Activity Aggregation - Establishment of relationships among diverse entities in time to identify meaningful events or activities.
3. Contextual Interpretation/Fusion - Analysis of data with respect to the context of the evolving situation including weather, terrain, sea-state, or underwater conditions, enemy doctrine, and socio-political considerations.
4. Multi-perspective Assessment - Analysis of data with respect to three perspectives: 1) the blue (friendly) force; 2) the red (enemy) force; and 3) the white (neutral) - how the environment affects red and blue perspectives.

Other authors [5-7] have come up with more detailed definitions of SA derived from the JDL model and linked with a conceptual model and functional decomposition in order to provide a better understanding of the concept. The commonality of each of these definitions coupled with research work at DREV, will be used to state the definition of SA that underlies the development of the generic model for STA presented in the next section.

Threat Assessment

TA is focused on the details necessary for decision makers to reach conclusions about how to position and commit the friendly forces. It can be viewed as a longer term diagnosis function to determine problems in the current situation and identify opportunities for taking corrective actions.

By coupling the products of SA with the information provided by a variety of technical and doctrinal databases, TA develops and interprets a threat oriented perspective of the data to estimate enemy capabilities and lethality, identify threat opportunities, in terms of the ability of own force to engage the enemy effectively, estimate enemy intent (i.e., provide indications and warnings of enemy intentions), and determine levels of risk and danger.

Hence, TA uses the situation picture from level 2 and what is known about the enemy doctrine and objectives to predict the strengths and vulnerabilities for the threat and friendly forces. In addition, the friendly mission and specific options available to the decision makers are tested within these strengths and vulnerabilities to guide decision making.

Key TA functions include: estimate enemy forces capability, predict enemy intent, identify threat opportunities, multi-perspective assessment, and offensive/defensive analysis.

REQUIREMENTS OF SITUATION AND THREAT ASSESSMENT

Similar to decision support system (DSS), requirements for situation and threat assessment (STA) are derived from characteristics of the threat, its environment, command and control (C2) operational requirements and C2 deficiency capabilities.

STA tasks, which can be characterized as either reactive (immediate response) or deliberative, are performed by either the human, the machine (computers, sensors, etc.), or a combination of both. All types of tasks must be designed and developed to manage the uncertainty and incompleteness of the perceived and processed data.

Under time-critical conditions[i], automation is essential leaving the operator to a passive role[ii]. In this context, STA reactive tasks performed by the machine must be real-time efficient. When the tactical situation permits deliberation, STA deliberative tasks must be designed and developed to perform synergistically with the operator. This suggests that the design and development of STA algorithms support the strengths and complement the weaknesses of operators by matching their perceptual and cognitive resources to the demands of the environment. For instance, taking advantage of the inductive intelligence of the human and deductive precision of the computer [8] could lead to a joint human-machine system for the interpretation of complex tactical situation where the human is responsible to generate hypotheses and the machine responsible for the validation of these hypotheses against data [9].

Another general requirement for STA is to support and to lighten the workload of the human. This could be done through the automation of some simple deliberative tasks (i.e., commercial flight correlation) or by monitoring and aiding the combat operator during the execution of standard operational procedures in engagement situation. STA enhancements have an impact on the interaction between the human-machine and modify the function allocation between them. For that reason, STA design and development requires taking into consideration the cognitive aspects of human information processing.

TOWARDS A GENERIC MODEL FOR SITUATION AND THREAT ASSESSMENT

Previous work on Data Fusion (DF) has tended to restrict humans to an observer's role and thereby exclude them from the process. In fact, the Joint Director of Laboratories (JDL) high-level functional model of the DF process suggests a hierarchy of sub-processes that could lead to the design and implementation of a totally automated system.

Our approach departs from this perspective. Rather, it is concerned with the reinsertion of the human in the loop by understanding the human's mental processes in order to optimally reallocate the functions between human and machine. Currently, in the operations room (OR), most of the reactive processing is done by the machine and the deliberative processing is done by the human. This is no surprise since the machine, with its computational power, is without doubt the most efficient means to execute reactive tasks. Similarly, the human with his reasoning capabilities is best suited for performing deliberative tasks.

In the context of Above-Water Warfare (AWW), Situation and Threat Assessment (STA) are active processes by which the decision maker in the ship's OR achieves awareness of the tactical situation in light of their goals. Endsley [10] defines situation awareness as the perception of the elements in the environment within a volume of time and space, the comprehension of their meaning, and the projection of their status in the near future. It can be interpreted as the operator's mental model of all pertinent aspects of the environment (process, state, relationships). Fig. 2 represents the three levels of situation awareness (the state) derived as products of the processes of STA.

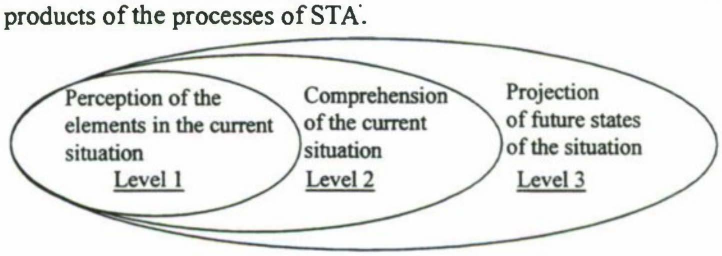


Figure 2. - Three Levels of Situational Awareness According to Endsley's Model

Obviously, STA is composed of two linked processes : Situation Assessment (SA) and Threat Assessment (TA). The SA process monitors the external environment to produce a situation description. Then, SA develops a higher level interpretation of the evolving dynamic situation description, based on a priori knowledge and transient information, and in terms of the current relationships among the perceived domain elements in the context of the operational environment and current mission goals. Therefore, the ultimate goal of SA is to determine the probable situation explaining the presence and the status of the observed entities in the environment in order to enhance awareness of the tactical situation. The result of SA is a coherent composite tactical picture of the current situation along with a short term prediction of the situation. The tactical picture is described in terms of groups or organizations of objects to be used for the

enhancement of the Commanding Officer's (CO) situation awareness and for the threat evaluation which is carried out in TA.

The second process of STA and TA evaluates and ranks threats on the basis of information obtained from dynamic tactical picture as well as from a priori knowledge and transient information. The result of TA is a ranked threat list used by Resource Management (RM) where decisions are made about how to use war fighting assets in support of the mission.

Efficient automation of some STA deliberative tasks should be carried out taking into consideration the cognitive aspects of human information processing. We propose a high-level functional decomposition of the processes of STA. We refer to this decomposition as the generic model for STA. It is motivated by Endsley's definition of situation awareness while, at the same time, giving consideration to the human-computer analogy that forms the basis of human information processing theory [11].

Although a difficult problem, the real challenge in command decision support technology remains in the implementation of a STA system that will complement or support decision-makers. Such a system must be designed so that the operator will trust and use it. The ultimate goal should be to implement a system that executes the STA process faster and more efficiently by a human-machine combination working in synergy.

High-level Functional Decomposition of STA

A first cut of the generic model for STA is illustrated in Fig. 3, focusing on a high-level functional decomposition of STA. The proposed model consists of a Perception Refinement module, a Threat Refinement module, a Situation Interpretation module, a Situation Projection module, a Monitoring module, and a Diagnosis module. The generic model has access to a priori knowledge and transient information and its behavior is modulated by a Meta-Controller taking account of processing priorities, processing time, and information quality .

The proposed model evolved from a three level descriptive model of situation awareness, where the first level is related to perception of the elements in the current situation, the second level is about comprehension of the current situation and the last level deals with projection of future situation states. This model yields a high-level functional decomposition of a multilevel STA process.

The inputs and outputs of the model

Inputs

The inputs of the generic model for STA are dynamic since they are provided by RM, Multi-source Data Fusion (MSDF), and external sources and they evolve in time. Inputs of the model can be categorized as organic and non-organic information.

Organic information

Inputs of the model controlled, collected and managed by assets under the CO's direct control can be defined or characterized as organic information (Datalink, Electronic Support Measures system, Identification Friend or Foe system). Organic information must be sufficiently timely and accurate to be used in real-time, responsive systems. Consequently, it can be used to produce a local tactical picture describing the situation in order to support all of the commander's activities at sea.

For instance, track information, as explained earlier, is provided by MSDF and constitutes the main part of the input to STA. It is characterized as organic information and consists of kinematic information (position, velocity, etc.) and the identity of a perceived entity. This information is sent to the Perception Refinement module. The RM process also provides the Perception Refinement module with track information. The track information provided is about own ship countermeasure actions (Fire-Control Solution) so that it can recognize the track and associate contacts pertaining to the own ship missile, which are perceived within the MSDF process. Here are some other examples of organic information:

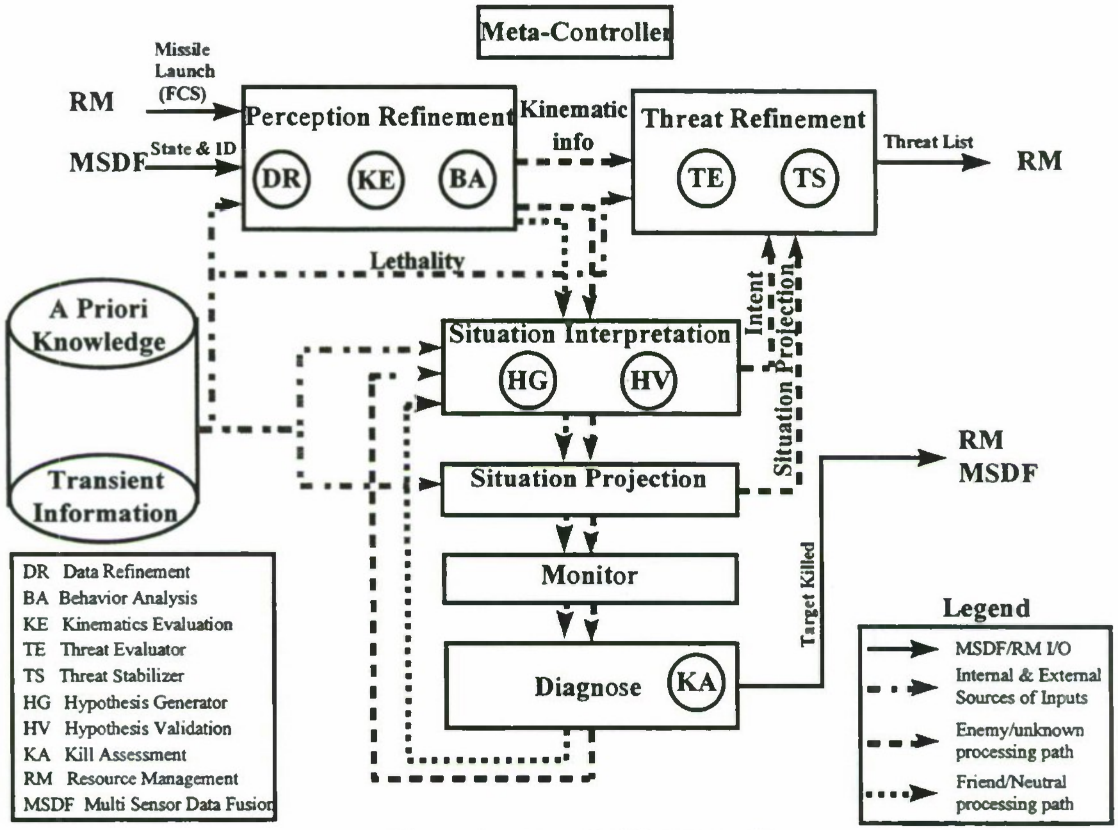


Figure 3. - Generic Model for STA

Non-Organic information

Inputs of the model collected by agents not under the CO's direct control, are referred to as non-organic information. This type of information is primarily used by ashore systems to provide some sort of global situational awareness. Some of this information can be of considerable interest to refine the local tactical picture. Non-organic information is less timely, reduced in accuracy, differently structured, and has differing identification confidence levels. For these reasons, it cannot be easily integrated into real-time and responsive systems.

Transient information such as intelligence reports or participating unit's information is non-organic and provides additional means for enhancing the commander's situation awareness. This information is used by the Perception Refinement module, the Threat Refinement module, the Situation Interpretation module, and the Situation Projection module. The information is used to refine track information, to generate cues for the Situation Interpretation module, to validate hypotheses, and to influence threat evaluation.

Other inputs

Another source of input to the model, which is not shown in Fig. 3, is the human contribution to the STA process through the Human Computer Interface (HCI). A complete and efficient automation of all deliberative tasks carried out in STA is not in the foreseeable future. So it is fair to say that these tasks are to be accomplished by a combination human/machine (to be determined) for efficient results. Consequently, it is obvious that humans will participate actively by providing inputs, based on experience, training and awareness of the situation within the STA process. The precise nature of this HCI remains to be determined.

Outputs

The results, or outputs, of the STA process consist of a stabilized and ranked threat list, result of Kill Assessment (KA) process, and feedback to the human concerning the tactical picture.

The threat list, which is obtained from the Threat Refinement process, is passed to the RM process, where decisions are made about how to use war fighting assets in response to threats.

The second output of the model, resulting from the KA process, is passed to both MSDF and RM processes, respectively, as a means of updating the tactical picture and providing input to planning and engagement actions.

Another output of the model, which is not shown in Fig. 3, is a high level interpretation of the tactical situation which is sent to the CO, through the HCI, in order to enhance the CO's situational awareness of the battle environment.

One has to wonder whether the Perception Refinement module, the Situation Interpretation module, and the Diagnosis module, described later in this section, should be producing outputs that could be useful to MSDF processing, e.g., hypotheses of possible clusters that could provide evidence of missing tracks.

A Priori Knowledge

A Priori Knowledge contains static information as a means to support the various processes providing the commander with a gain in a level of situation awareness. This knowledge is a component within the model as opposed to inputs derived from external sources. A Priori Knowledge is used by the Perception Refinement module, the Threat Refinement module, the Situation Interpretation module, and the Situation Projection module. The information is used to refine track information, to generate cues for the Situation Interpretation module, to validate hypotheses, and to influence threat evaluation. A priori knowledge can be mapped in a human information processing model as long term memory. Here are some examples of a priori knowledge sources:

- social and political
- geographical
- platforms characteristics
- mission guidelines
- weapons characteristics
- corridor and flight paths
- lethality
- emitter characteristics
- doctrines
- etc...

The Perception Refinement module

The Perception Refinement module corresponds to the first level of Endsley's model. All of the processing related to low-level information such as entities (track data) and groups of entities (clusters) is addressed in this module. The first goal of this module is to refine data by examining track attributes (position, identity) from MSDF for incompleteness and contradictions, and then attempting to establish relationships among these entities in order to form clusters. The second goal of the Perception Refinement module is to estimate kinematic parameters for weapon engageability calculations which are performed in RM. The last goal of this module is to perform behavior analysis of entities and/or clusters in order to help refine the data set and also to provide the necessary requirements (cues) for the interpretation and understanding of the tactical situation done within the Situation Interpretation module (explained in a later subsection).

If the allegiance of the track, as first estimated by MSDF and refined in this module, is either a foe or an unknown, the refined data is passed directly to the Threat Refinement module defined in a later section. If time permits deliberation, the process will, independently of the track's allegiance, transmit to the Situation Interpretation module the cues generated by the Behavior Analysis sub-process for higher level situation awareness processing.

The Perception Refinement module is composed of three sub-processes: the Data Refinement (DR) sub-process, the Kinematic Estimation (KE) sub-process and the Behavior Analysis (BA) sub-process.

The idea of DR is to refine the track data (position, identity) already generated by MSDF by examining the data set for incompleteness and contradictions and to establish relationships among the entities (in terms of proximity, functionality, and dependency) with the help of external data sources if necessary. In the process, no inferences about the situation are generated. The only results obtained from this sub-process are perceptual refinements.

The KE sub-process is used to compute kinematic information (mean line of advance, closest point of approach (CPA), time of flight (TOF), etc.) of a track in preparation for Threat Refinement processing, and for weapon engageability calculations done in RM. A history function, which records the positional

tracking and identification information in time, is needed to accomplish these kinematic calculations. The recording of the track information will also allow us, through a history function, to address enemy information countermeasures (i.e., information warfare).

BA is the last sub-process under Perception Refinement and is used to analyze the behavior of entities and/or clusters in order to help refine the data set. Also, BA becomes a preliminary step for the second level of the situation awareness model by providing processing cues, or evidence, about track behavior, or cluster status, to the Situation Interpretation module for the interpretation and understanding of the current tactical situation. The cues are obtained based on DR analysis of refined information (track and cluster's kinematic data and identification), from a priori knowledge and from transient information (i.e., electronic emissions from Electromagnetic Sensor Management (ESM), datalink, information from participating units, etc.). BA includes functions, such as corridor correlation, maneuver/pattern identification which generate the required cues to make inferences about the tactical picture. The output of BA, if time permits deliberation, is interpreted within the Situation Interpretation module which is explained below.

The Situation Interpretation module

The Situation Interpretation module is the final processing step to achieve the second level of situation awareness. The Situation Interpretation module is defined as a deliberative module that generates and validates hypotheses about the current tactical situation based on the outputs of the Perception Refinement module, a priori knowledge and transient information. Therefore, the Situation Interpretation module explains the presence of the perceived entities and determines the intent of enemy or unknown tracks. In addition, if the track's allegiance is currently perceived as a foe or unknown, the module influences the threat assessment done within the Threat Refinement module as explained later in this chapter. Finally, the results of this module are passed to the Situation Projection module, independently of the track's allegiance.

The Situation Interpretation module is composed of two sub-processes: the Hypothesis Generator (HG) and the Hypothesis Validation (HV). The first is used for each new entity or event to generate one or several hypotheses about the probable situation causing the perceived domain element (track, cluster). These new hypotheses are obtained based on the outputs of the Perception Refinement module, a priori knowledge, and transient information and are validated by HV for inconsistencies, conflicts of information, and potential inaccuracies due to incomplete data.

The goal of the HG process is to provide possible explanation of the role and purpose of each perceived entity within the domain. In the case of friend/neutral allegiance, the intent is relatively straightforward to determine due to the cooperative nature of the entity. Hypotheses about the enemy/unknown's intent are likely to be more variable due to the non-cooperative nature of the entity. Independently of the allegiance, all hypotheses need an iterative validation process to acquire and modify the confidence level associated with each hypothesis. To do this, the HV process takes cues from the Diagnose module (explained later in a subsection) along with an updated situation description and modifies the hypotheses and their confidence levels to reflect current understanding of the state of the world. Finally, the resulting hypotheses are fed into the Situation Projection module for further processing and also into the Threat Refinement module for threat evaluation calculations.

The Situation Projection module

The Situation Projection module concerns the last of the three levels of the Situation Awareness model: projection of future states. The Situation Projection module is a deliberative process that generates hypotheses about the future states of the tactical situation based on outputs of the Situation Interpretation module and the a priori knowledge. The results of the Situation Projection module are passed to the Threat Refinement module as input to threat number calculations and then submitted to the Monitor module to continue the SA processing.

The Monitor module

The Monitoring module is in fact a process that stores expectancies such as hypotheses about future events and anticipation of KAs, and monitors the situation to collect cues until a diagnosis can be given which is completed by the Diagnose module. The goal of this module is to monitor the situation for potential violations of expectancies generated by the Situation Projection module, which suggest that the current situation interpretation is in error.

The Diagnose module

The Diagnose module measures the discrepancies between expectancies and the currently perceived state of the world and diagnoses the nature of these discrepancies. These diagnoses are fed back to the Situation Interpretation module for further interpretation and validation.

The Diagnose module also assesses expectancies related to countermeasures taken by RM, called Kill Assessment Briefly, KA consists in assessing the kill (soft or hard) of an entity by monitoring the results of own ship countermeasure actions. The result of this particular expectancy is sent back to MSDF for an update of the tactical picture and to RM for an update of engagement plans.

The Threat Refinement module

This module assesses potential threats and produces a stabilized and ranked threat list based on the opportunity, lethality, and intent of the threat, and a short term prediction of the situation. The Threat Refinement module is composed of two processes: the Threat Evaluator (TE) and the Threat Stabilizer (TS).

The TE process evaluates threat of unknown, or enemy tracks, based on opportunity and lethality information. The threat assessment is also refined using intent and situation prediction outputs from the situation interpretation and situation projection processes respectively.

The opportunity can be determined with respect to own ship or another ship. For instance, the opportunity of air threats is determined, on the Canadian Patrol Frigate (CPF), from the closest point of approach (CPA), and time of flight (TOF) calculations along with the target's mean line of advance and velocity.

The lethality estimation is based on a priori knowledge and transient information (intelligence reports for example) of the target characteristics, the weapons on board the target, the characteristics of these weapons, and their status.

The intent of an air threat is estimated in the situation interpretation module and can raise or lower the threat level depending on current track behavior. For instance, if an anti-ship missile is locked onto a ship and the ship has proof of this fact from its sensors, this situation is obviously more dangerous than when the missile seeker head is passive. As another example, a multi-mission bombing aircraft that is only doing reconnaissance would be considered less dangerous than the same aircraft if it were carrying out a bombing mission. In the latter case, therefore, the intent would influence the threat assessment.

Finally, a situation prediction used together with the appropriate subfunctions of the diagnosis process could give better estimates of opportunity and intent and thus refine the threat assessment.

The TS is a process that prevents the outputs of TE from oscillating. In the case of maneuvering targets approaching a warship, the CPA can vary between large positive values and almost zero values. This variation causes instability in any list of absolute and relative threat levels. Once the list of threats has been stabilized, TS ranks them through the use of a prioritization function whose results are passed to RM.

An important issue remains to be addressed. Should stabilization of the outputs of TE (threat values) take place within the Threat Refinement module or at their origin (origin of the inputs of TE)? Future work on the Generic Model for STA should resolve this issue.

The Meta-controller

The information processing in STA plays a key role in the decision making concerning use of the ship's weapon and sensor systems in the AWW. However, in a manner analogous to human information processing itself, it needs to be controlled and regulated as part of a generally reflective executive function. This functional capability in the generic model is provided by the Meta-controller.

The requirement for a Meta-controller can be traced to a number of critical characteristics of the AWW environment which is highly dynamic. Critical events happen at indeterminate times, and at any given moment multiple contacts may be under investigation, assessment and evaluation, at various stages in their processing chain. There can be time-varying priorities on processing a particular contact, depending on the perceived risk to the achievement of mission goals posed by the contact. The processing itself needs to account for the time pressure for producing its results, for example, by limiting the amount of information used in doing the processing. The information available in the AWW may be ambiguous, incomplete, erroneous, or imprecise. Information quality therefore needs to be monitored for its effect on the amount and nature of the processing required. Initiating actions (involving sensor management, navigation maneuvers, etc.) to acquire additional information in support of further STA processing of a

contact may also be part of the Meta-controller's executive strategy in cases of inadequate information on that contact.

The Meta-controller, therefore, provides the necessary flexibility for monitoring the status of STA processing and opportunistically controlling this processing so as to respond to the various data-driven and goal-driven demands on this processing as events in AWW unfold.

Limitations of the Model

STA is a complex process which deals with issues such as mission assessment, battle damage assessment, multi-perspective assessment, plan monitoring, and risk assessment to list a few. Here, we have shown a first cut of a model oriented mostly on one aspect of the STA process, which is engagement situation, but with considerations for human's situational awareness.

The generic model only illustrated a philosophy for building a system that assesses a tactical situation and its related threats. It is based on an attempt to understand the human's cognitive processes that derive mental representation (or situation awareness) of the state of the world.

The methodology for building this model is not addressed in this document. A requirements definition is needed to address issues, such as decision-making requirements, task/function identification, and automation at a generic level, through a top-down analysis of the problem. Toward the validation, refinement, and development of the generic model, knowledge of the human's cognitive processing is essential and is at the heart of this approach. Expertise in the domain of cognitive science will be needed to look at issues such as tradeoffs in resources, task/function frameworks consistent with the multilevel cognitive representation of Situation Awareness, the control framework for invoking these functions (goal driven vs data driven), and mechanisms for the meta-control of STA operations. These research activities are essential to validate the generic model and lead to its eventual implementation.

Other research topics such as information warfare and a model for handling imperfect information in the STA process are important issues not addressed by this generic model and which need to be investigated to enhance or refine the generic model.

Implementation of Baseline based on the Generic Model

The expert system underlying the Command and Control (C2) System must meet several requirements which include a knowledge-based implementation [12] real-time efficiency, distributed (multiprocessor) environment, design flexibility, and guaranteed real-time execution for the multi-source/situation and threat assessment/resource management (MSDF/STA/RM) components. There is currently no commercial knowledge-base system (KBS) on the market which can fulfill this set of requirements and, consequently, a more powerful KBS shell needs to be designed as part of this project.

The above requirements point in the direction of a specific class of expert systems, namely the blackboard-based problem-solving engines, or Blackboard Systems. The blackboard architectural model of problem solving has been developed for the purpose of speech analysis, yielding Hearsay-I system which has been further developed into the classical blackboard architecture with Hearsay-II. More formally, the blackboard model of problem solving is a highly structured special case of opportunistic problem solving, in which pieces of knowledge are applied either forward (i.e., data-driven) or backward (i.e., goal-driven) at the most "opportune" time in order to construct a solution to the problem.

These systems provide a powerful architecture to deal with problems whose nature is such that 1) many diverse, specialized knowledge representations are required; 2) heterogeneous problem-solving mechanisms need to be integrated under a common framework; 3) knowledge is often uncertain or incomplete; 4) the application is to be developed by numerous developers; and 5) problem-solving activities require multilevel reasoning or flexible, dynamic control mechanisms. Moreover, the Blackboard System architecture offers a great deal of flexibility and adaptability, supports incremental and structural problem-solving strategies, and is well suited to find solutions efficiently in a large search space. Clearly, those benefits are needed to a large extent in a real- time environment. Blackboard systems are usually not suited for simple problems and, therefore, their use should be limited to complex applications. In the case of a Canadian Patrol Frigate (CPF) Command and Control System (CCS) integrating MSDF/STA/RM functionality there is no doubt that this criteria is easily met.

All entities in our KBS shell implementation are represented as data objects. These data objects are connected together by various semantic relations (has_member, is_part_of, can_produce, etc.). Since the relations among data objects may vary depending on the contextual environment (e.g., a plane may have a

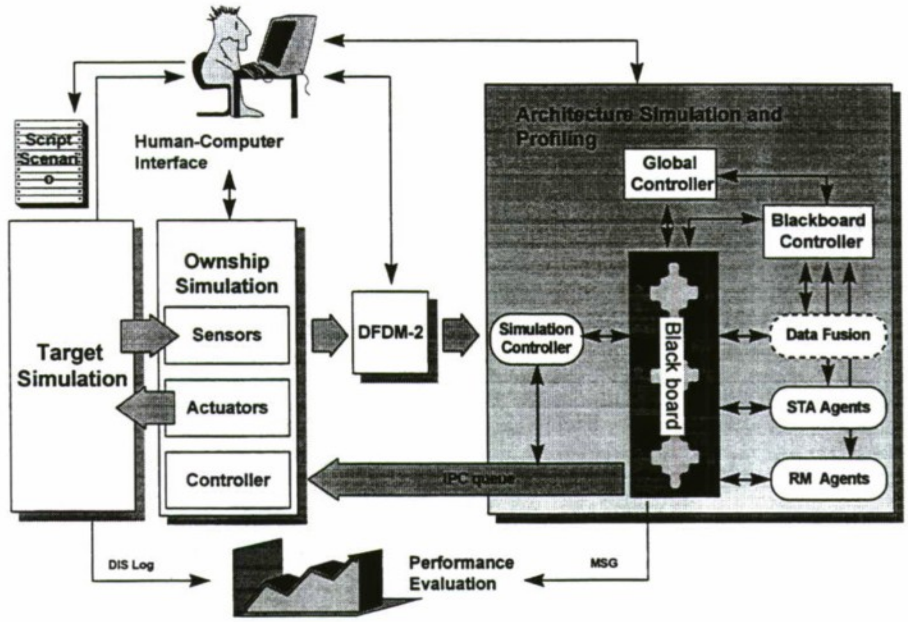


Figure 4. - Closed-loop Environment and Blackboard System Architecture

different meaning in a warfare context or in the context of a child's toy), special functions called context functions are introduced to select or prune these relations given a specific context. Data objects first need to be instantiated on the blackboard — this process is called data instantiation. Once on the blackboard, the data can be in an active or passive state. Only the active data can trigger the agents which reside in the various knowledge-based sources. Similarly, these agents can also be active, or passive, and only the active agents whose input data are active will be executed. It is the task of the blackboard controller to look for pairs of active agent/data and execute them. The agent execution is, thus, controlled by activating, or de-activating, the agents and the corresponding input data. During the execution of the agent function, the data is accessed directly on the blackboard, processed, and put back on the blackboard by the agent (or deleted). New data objects can also be instantiated on the blackboard by these agents. While most commercial expert systems available on the market rely on interpreters (i.e., instructions are interpreted sequentially), our KBS shell is fully compiled in C++, a feature which increases the execution speed by a factor of 100 to 1000. Currently, more than 10,000 agents can be activated and executed in less than 0.4 second on a single-processor Ultra Sparc workstation running under Solaris 2.5.

For the Basic Demonstration Model, we have designed a set of 16 agents to perform the Threat Evaluation/Weapon allocation (TE/WA) CPF functionality. These include the STA agents: opportunity calculation based on closest-point-of-approach (CPA) and time-of-flight (TOF), air safety corridor correlation, threat evaluation, threat ranking, kill assessment (KA); and the RM agents: point-to-point engageability, target/weapon pairing, resource assignment, and weapon order generator. Target and own ship simulations are provided by a software called Computer-Aided Software Engineering- Advanced Testing Technologies, Inc. (CASE-ATTI). Scenarios are predefined and read from script files. The own ship simulator, which simulates the sensor reports, sends the data to the Data Fusion (DF) component of the system (here, Data Fusion Data Model 2 (DFDM2)), which in turn produces fused tracks on the blackboard. These tracks are then processed by the STA and RM agents to produce a prioritized threat list as well as weapon recommendations and weapon orders. If the system decides to launch a surface-to-air missile towards an incoming enemy target, a new track corresponding to the missile is being created and fed back into the simulation environment. These missiles can then interact with the incoming targets and eventually destroy them. These elements, thus, constitute a unique closed-loop environment, illustrated in Fig. 4, in which MSDF, STA, or RM algorithms can be tested.

Several scenarios have been built to validate the algorithms implemented so far, and also to evaluate the speed of the KBS shell. In one scenario, four targets are approaching the own ship from various directions at different velocities. As the targets enter the quick reaction (QR) range of the own ship, surface-to-air missiles are automatically assigned, launched, and all enemy targets are successfully intercepted. This scenario spans 800 seconds of simulation time and it is executed in about 80 seconds on an Ultra Sparc. It is worth pointing out that in this 80 second period, only 2 seconds are used by the KBS to process the 2200 individual contacts, and a total of more than 35,000 STA and RM agents are actually being fired; the rest of the time is used by the environment/own ship simulator and data fusion module. Similarly, in another 800 second scenario executed in 240 seconds and involving 14 incoming targets (2 of which are missiles launched from air platforms), the KBS shell processes 6600 contacts — executing more than 100,000 agents — in about 7 seconds. Even though the current functionality of the STA and RM agents is minimal, these figures illustrate, nevertheless, that the KBS shell itself (e.g., data instantiation, data activation, agent execution, controller, etc.) represents only a minor contribution to the total execution time of our system.

CONCLUSION

This document presented a generic model for Situation and Threat Assessment (STA). The proposed model is currently being used in an exploration of real-time issues for an integrated multi-source/situation and threat assessment/resource management (MSDF/STA/RM) system for the Canadian Control Frigate (CPF).

Unlike the traditional model of STA based on the level of abstraction of the data, the philosophy of the proposed model is concerned with the reinsertion of the human-in-the-loop by taking into consideration the human's mental processes that lead to the development of his situation awareness. The generic model for STA evolved from the three level Situation Awareness model of Endsley, knowledge and models of the human's cognitive processes, leading to a high-level functional decomposition of a multilevel STA process. This feature of the model leads to efficient automation of some deliberative tasks within STA and currently done by the human. In addition, this model permits the human to contribute actively to the STA process through the Human Computer Interface (HCI). Therefore, the implementation of the proposed model is being pursued as part of the design of a real-time decision support system for combat system operators of the CPF.

A baseline of the generic model for STA is currently being implemented and integrated with baselines of MSDF and RM. The results of the implementation will be presented in a future document.

The results of this research are expected to contribute to the Defence Research Establishment Valcartier's (DREV's) investigations of enhancements to the CPF's Command and Control System (CCS) as part of the mid-life upgrade of the CPF expected early in the next century.

Future Work

A subset of all the functionality within the generic model, called a baseline for STA, is currently being implemented through a collaborative activity between Lockheed Martin Canada and the DREV. The complete implementation of the baseline is scheduled for 1998.

Short term work will refine our understanding of the human's cognitive processes for achieving a situational awareness by conducting a top-down analysis of the domain of command and control (C2) in order to define the human's decision requirements according to a established cognitive engineering methodology.

Medium term work will consist mainly in investigations and study to acquire expertise in the area of cognitive science, situation awareness, information warfare, and uncertainty management. The expertise gained is expected to yield an enhanced generic model for STA.

The short term and medium term work will be accomplished through local research and development investigations, collaborative efforts with university and industry, and through cooperative research with foreign countries via The Technical Cooperation Program (TTCP) and memorandum of understanding agreements.

ACKNOWLEDGMENT

The authors wish to thank our colleagues in the Data Fusion and Resource Management Group at Defence Research Establishment Valcartier and in the Research and Development Department at Lockheed Martin Canada for their invaluable discussions and comments about this model.

REFERENCES

1. Chalmers. B.A., Carling. R., Roy. J., Paradis. S. and Bossé. É., "Study of Real-Time Issues for an Integrated MSDF/STA/RM System for the CPF: A Detailed Description of SRTE Project", DREV-R-9629/97, January 97, UNCLASSIFIED.
2. Chalmers B.A., "Design Issues for a Decision Support System for a Modern Frigate", this Proceedings, 1997.
3. Llinas, J. and Waltz, E.L., Multi-Sensor Data Fusion, Artech House, Norwood, MA, 1990.
4. DFDS Panel, "Functional Description of the Data Fusion Process", Data Fusion Development Strategy Panel, Document Prepared for the Office of Naval Technology (ONT) Data Fusion Development Strategy, November 1991.
5. Roy, J., Chalmers, B., Carling, R.and Bossé, É, "Overview of Shipboard Data Fusion and Resource Management R&D Results and Rationale for its Real-Time Implementation in the ASCACT Testbed ", DREV-TM-9518/95, April 1996, UNCLASSIFIED.
6. Anthony, R.T., Principles of Data Fusion Automation, Artech House, Norwood, MA, 1995.
7. Llinas, J. and White, F.E., "Data Fusion : The Process of C3I", Defense Electronics, v.22, no.6, June 1990, pp. 77-83.
8. Ballas J.A., McFarlane D.C., Achille L.B., Stroup J.L. and Kushnier S.D., "Interfaces for Intelligent Control of Data Fusion Processing", Naval Research Laboratory, NRL/FR/5513-96-9806, April 1996.
9. Licklider J.C.R., "Man-computer Symbiosis", IEEE Trans. on Human Factors I, pp. 4-11, 1960
10. Endsley, M.R., "A Methodology for the Objective Measurement of Pilot Situation Awareness", Paper presented at the Aerospace Medical Panel Symposium, Copenhagen, Denmark, October 1989.
11. Wickens, C.D. and Flach, J.M., "Information Processing", in Human Factors in Aviation, E.L. Wiener and D.C. Nagel eds, Academic Press Inc., San Diego, pp. 111-155, 1988.
12. Ballard, D. and Rippy, L., "A Knowledge-Based Decision Aid for Enhanced Situational Awareness", 13th AIAA/IEEE Digital Avionics Systems Conference, Phoenix, Arizona, November 1994, pp.340-347.

Additional author information

S.P. : Phone : (418) 844-4374; Fax : (418) 844-4538; E-mail : stephane.paradis@drev.dnd.ca

B.C. : Phone : (418) 844-4698; Fax : (418) 844-4538; E-mail : bruce.chalmers@drev.dnd.ca

P.B. : Phone : (514) 340-8310 (ext. 8536); Fax : (514) 340-8318; E-mail : pierre.bergeron@lmco.com

[i] When a number of quick reaction criteria is met by the threat (i.e., speed, range from ship, Closest-point-of-approach...).

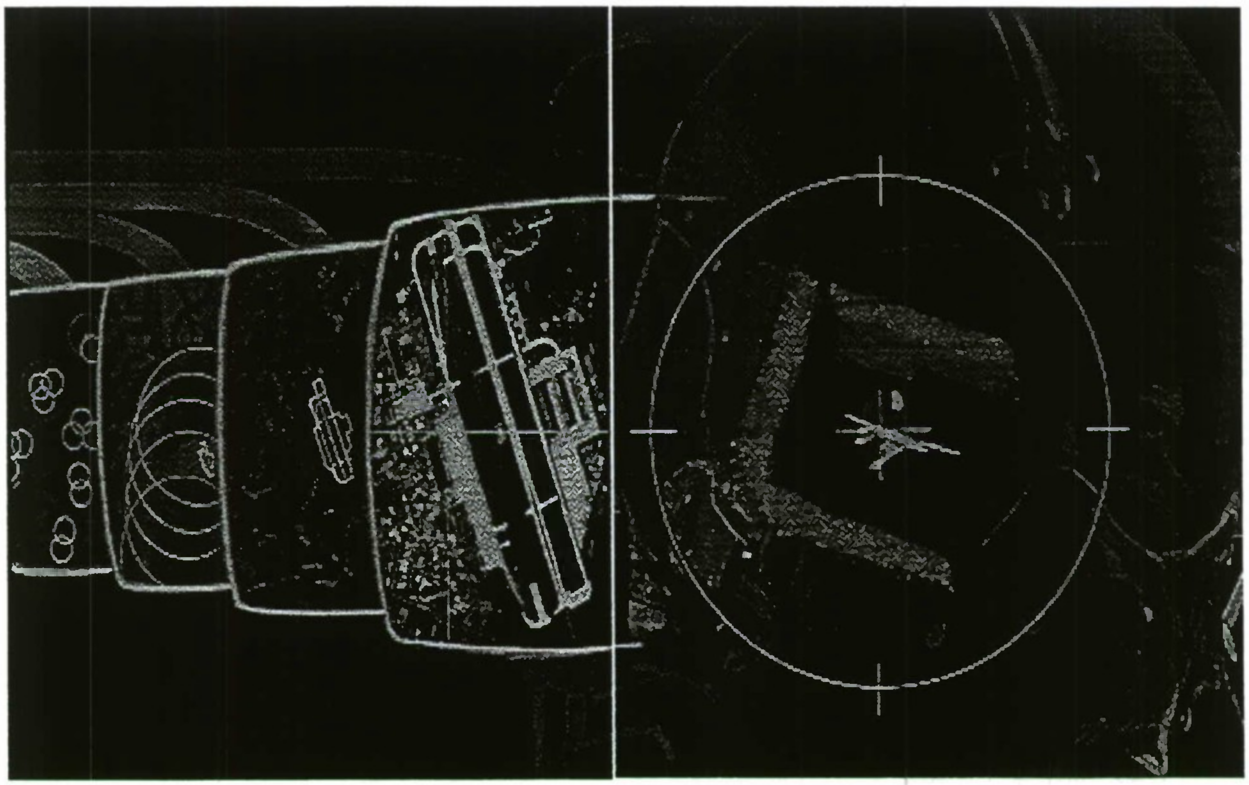
[ii] In an automatic mode, the operator acknowledges all actions to be taken. In an autonomous mode the operator is out of the loop.

AIRBORNE TACTICAL INFORMATION MANAGEMENT SYSTEM PROGRAM DESCRIPTION (ATIMS)

Randal Holt
VEDA Inc.
1755 Jefferson Davis Hwy., Suite 200
Arlington, VA

Mark Storer
GDE Systems, Inc.
P.O. Box 509008
San Diego, CA

Roger Logan
NAVAIRSYSCOM
(PMA-233)
Patuxent River, MD



INTRODUCTION The United States Navy Program Executive Office for Tactical Aircraft Programs, PEO(T), is developing an Airborne Tactical Information Management System (ATIMS) capability under the Space and Naval Warfare Systems Command (SPAWAR-133) Real-Time Support for Joint Power Projection (RTS/JPP) Core Technology program. The ATIMS program is leveraging modular processing, advanced display and virtual reality technology to demonstrate a capability that provides enhanced flight situation awareness of engagement parameters, exploitation of information for mission alternative selection and more responsive unit level mission planning and rehearsal. As illustrated in Figure 1, the ATIMS program is focused on developing tactical information management and cockpit automation technology to simultaneously reduce pilot workload, reduce system avionics cost, enhance operational flexibility and increase mission effectiveness for air-to-air or air-to-ground combat operations. In a cooperative effort with the Air Force Avionics Directorate, Wright Laboratory (WL/AART-3), the Real-Time Targeting Concept Development program is providing ATIMS concept development, program development and systems integration support.

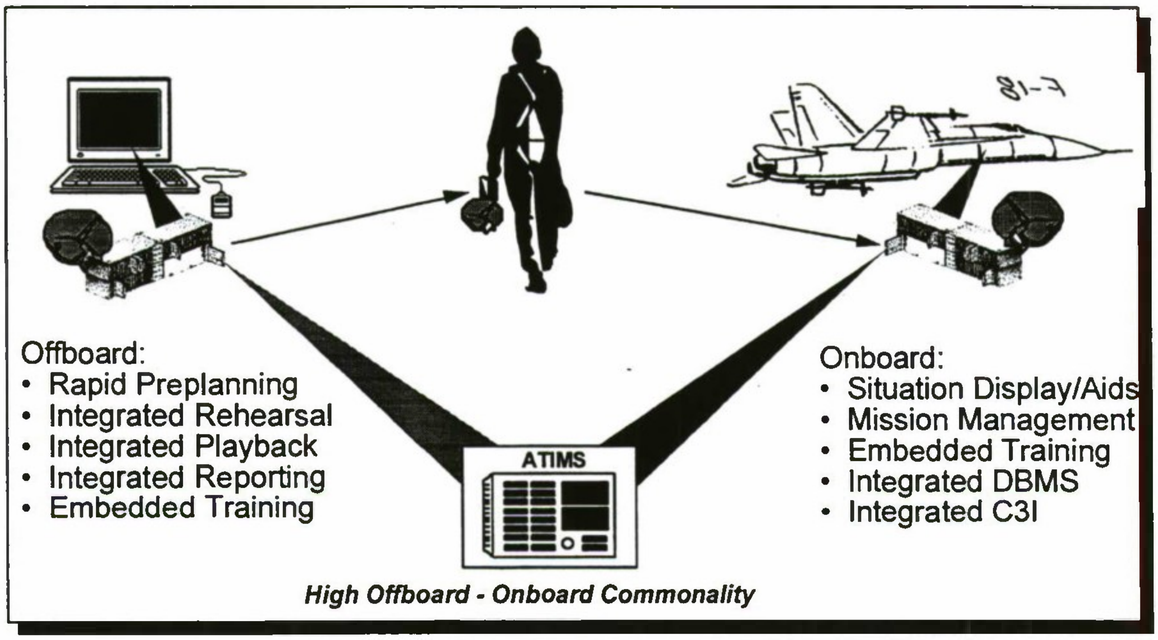


Figure 1. The ATIMS program is applying modular processing and open systems architecture technology to achieve high onboard/offboard functional commonality.

PROBLEM The ATIMS program is attempting to solve a dichotomy of competing trends - the rapid transition of new technology to meet increasing operational mission demands while simultaneously minimizing system/avionics (cost) impacts. Key challenges include:

- Reduced mission timelines.
- Exploitation of real-time information.
- Cockpit information overload.
- Increasing aircrew workload.
- Effective in-flight situation awareness.
- Inflexible infrastructure for insertion.
- System/avionics cost escalation trends.
- Inter and joint service interoperability.
- Legacy of stove-pipe avionics systems.

A fundamental driver behind ATIMS requirements is the larger need for reduced mission timelines to achieve higher effectiveness against time critical fixed and mobile targets (TCT's). The short

vulnerability cycle time window of TCT's requires rapid in-flight fixed target updates, unplanned threat avoidance replanning and adaptive end-game planning for mobile target prosecution.

As illustrated in Figure 2, ATIMS complements related initiatives such as the Real-Time Information Into-Out-of the Cockpit (RTIC/RTOC) programs which are evaluating methods to provide real-time information to shooter aircraft. ATIMS focuses on how to effectively exploit, manage and utilize offboard information as well as mitigate potential information overload. ATIMS is also concentrating on Command, Control, Communications, Computers, and Intelligence (C4I), planning and interoperability issues by addressing the stovepipe avionics systems across platforms as well as the lack of functional data continuity between onboard and offboard systems. A critical ATIMS concern is system life cycle cost which becomes acute when avionics cost trends are considered.

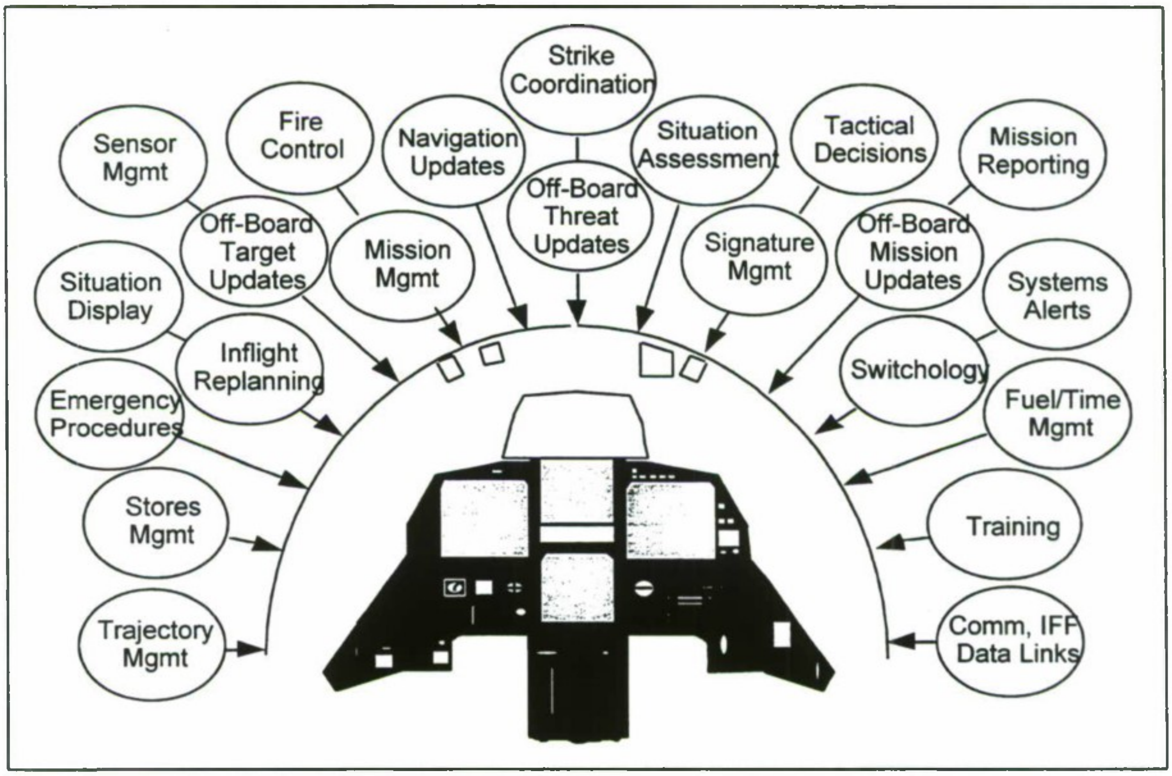


Figure 2. Key challenges the ATIMS program is addressing include cockpit information overload, increasing aircrew workload and effective exploitation of real-time information.

TECHNICAL APPROACH ATIMS is leveraging modular processing and virtual reality (synthetic environments) technology to demonstrate a capability that synthesizes information rapidly and presents an intuitive course of action for the aircrew.

From a system perspective, the ATIMS concept will allow a seamless flow of information between penetrating strike aircraft and follow-on strike aircraft. Integrated planning and preview operations will take advantage of perishable strike information through rapid C4I digital battlefield database updates and mission playback/analysis.

Figure 3 depicts the ATIMS circa 2000 concept vision which employs a modular mission management capability that supports offboard preflight and onboard inflight functions, including:

- Integrated unit level planning, training and rehearsal.
- Inflight mission management and mission replanning.
- Automated offboard C4I integration.
- Onboard/offboard sensor integration.
- Enhanced human computer interaction.

• All weather day/night mission visualization.
• Integrated tactical mission reporting and playback analysis.

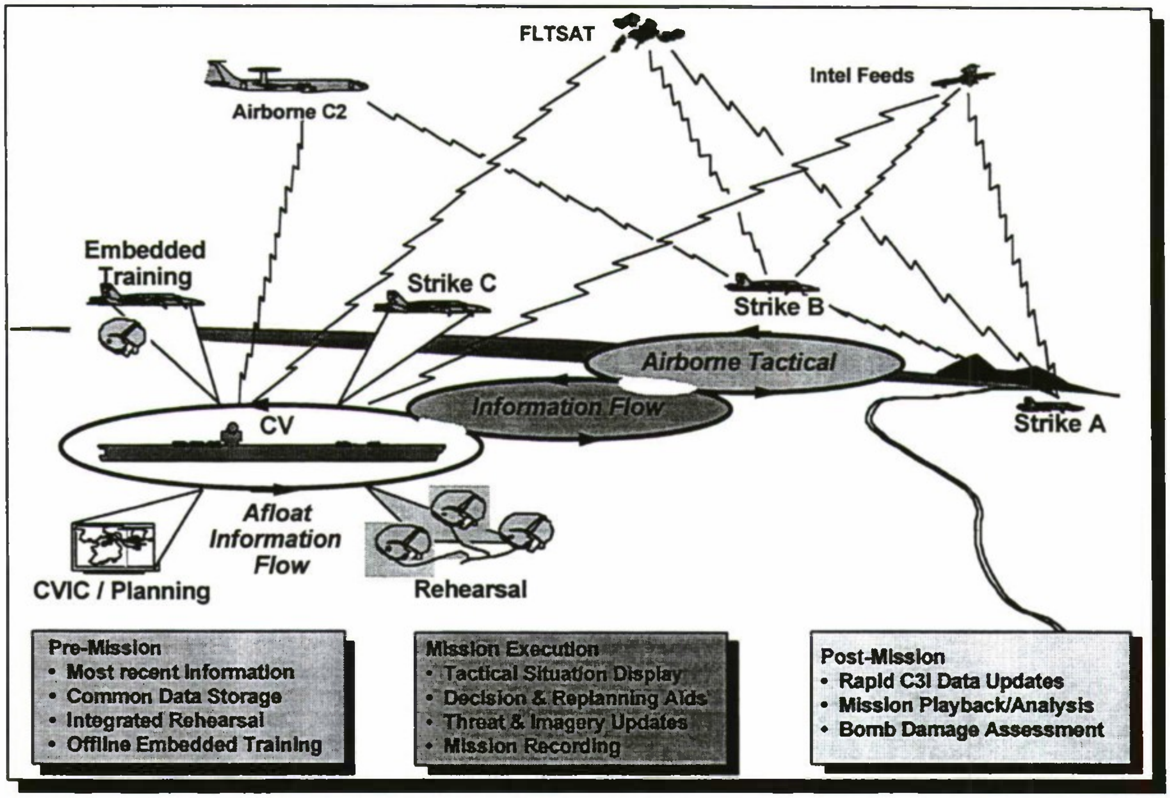


Figure 3. ATIMS is focused on developing a seamless flow of strike information to make effective use of perishable mission data to support sequential strike package readiness.

During ground-based pre-mission preparations, the aircrew would use ATIMS to interface with existing offboard system networks for access to mission planning functions and source materials (digital battlefield databases). During inflight operations (Figure 4), the aircrew would use the same ATIMS functionality to interface with the aircraft's avionics systems and helmet mounted displays. Sample information management functions include system status monitoring, flight trajectory management, weapon delivery automation, sensor management, tactical situation display, tactical decision aids, adaptive mission replanning and embedded training. Besides cockpit automation, this technology provides targeting flexibility by enabling tactical aircraft to effectively respond to offboard air tasking changes, theater intelligence updates, local onboard intelligence updates and unplanned mission events.

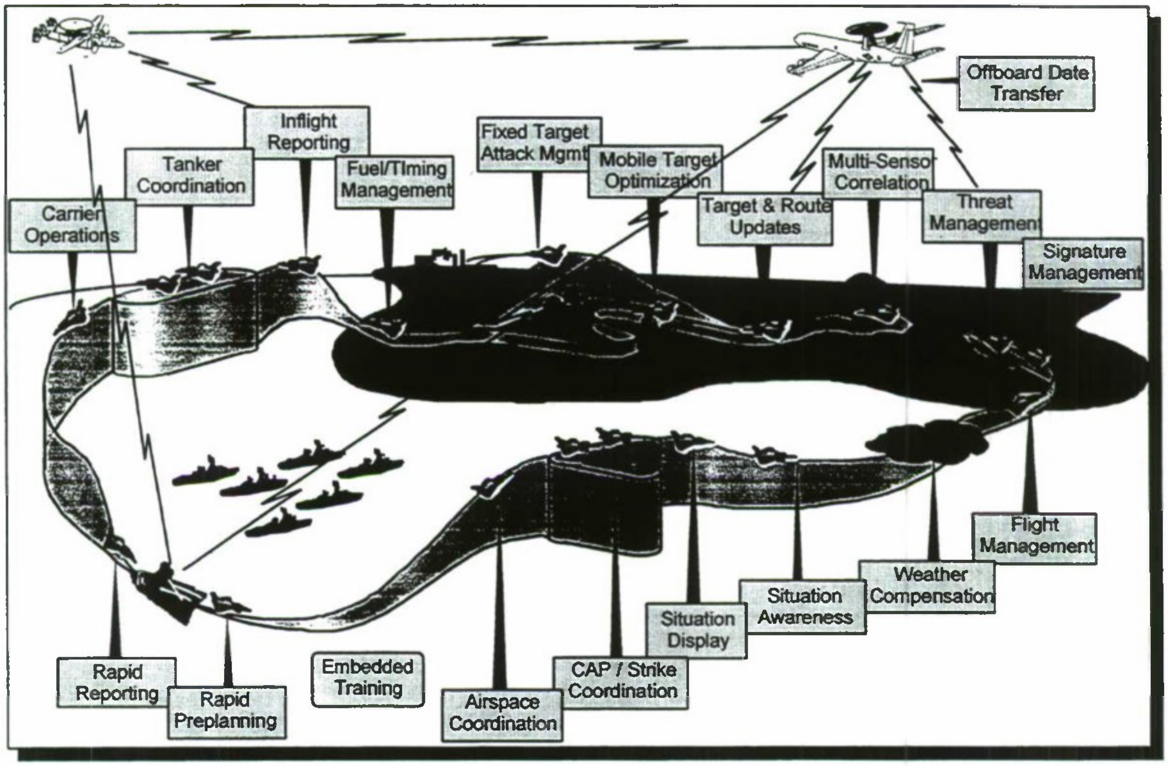


Figure 4. ATIMS is evaluating the partitioning of on/offboard functions to streamline system operations and enhance inflight operations against time critical targets.

Onboard information management requires many of the same functions as the offboard preplanning process, but in a real-time environment. Thus, software and hardware developed for the onboard capability is striving to embody the same characteristics as the offboard elements to ensure consistent behavior inflight and avoid software development and validation duplication.

By isolating the ATIMS components from the aircraft avionics, both multi-platform flexibility and lower life cycle cost is also achieved. Key technical development features include (Figure 5):

- Integration of new component technologies.
- Open systems/software architecture.
- Modular processing architecture.
- Integration of force, unit and intelligence information.
- On/offboard functional data commonality.
- Local and distributed connectivity between planning, rehearsal, simulation, training and inflight operations.

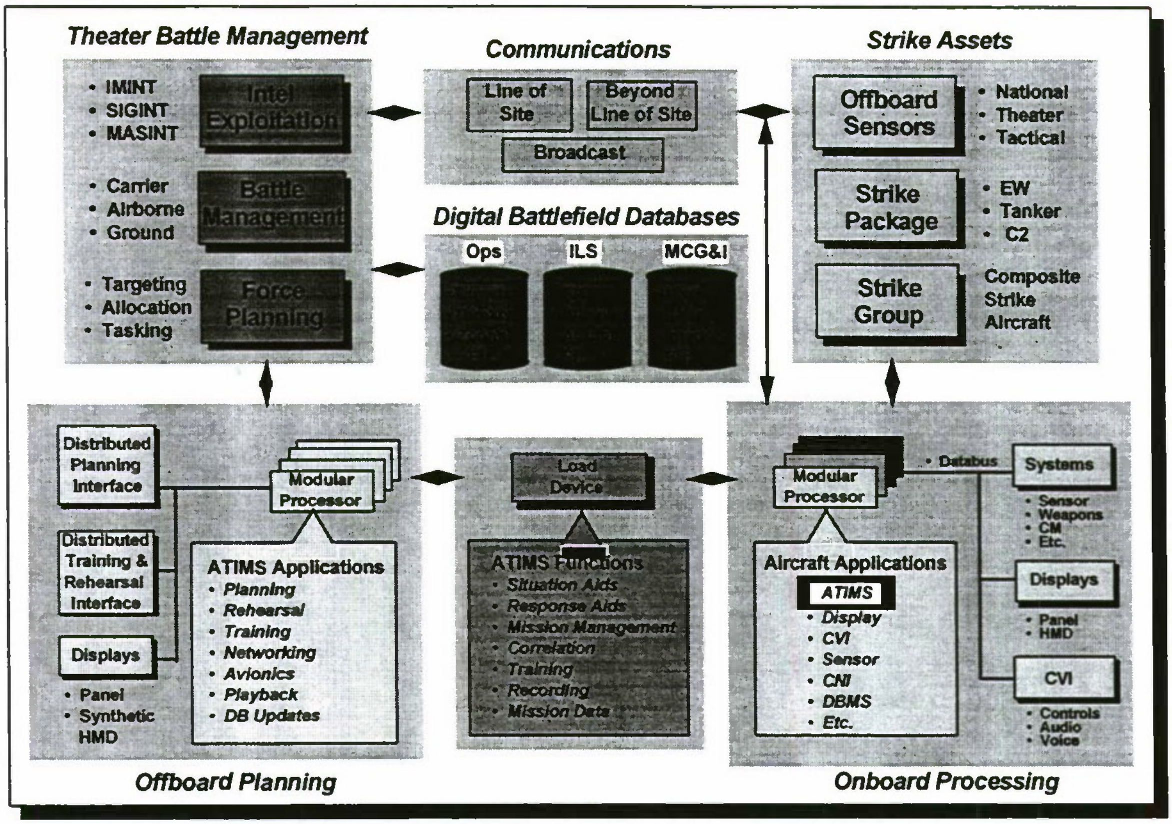


Figure 5. The ATIMS approach employs modular processing to permit high onboard/offboard software reuse and maintainability.

Key technologies being incorporated include:

- Advanced panel and helmet displays.
- High density electronics.
- Digital datalinks.
- Automated planning.
- Decision aids.
- Data compression.
- Synthetic environments.
- Aircrew interface techniques.
- Data visualization techniques.
- Multi-level security.
- Commercial open architectures.

DEVELOPMENT APPROACH ATIMS is one of four major components of the SPAWAR RTS/JPP Core Technology Program. Other components include the Scaleable High Performance LAN (SHPL), the Advanced Power Projection Planning and Execution System (APPEX), and the Tactical Situational Display - Joint Terminal (TASID-JT). The combined elements of RTS/JPP concentrate on different parts of information flow with the goal of producing tools that serve a mission management architecture.

The long term ATIMS plan embodies a four phased, multi-year strategy that includes integration of emerging technologies, periodic technology risk reduction laboratory and flight demonstrations and incremental technology insertion as follows:

- **Phase I** - Proof of concept via avionics hotbench simulation of limited in-cockpit mission replanning, offboard mission planning and playback and offboard threat data transfer as well as background long-term concept development studies.
- **Phase II** - High fidelity flight simulation of advanced tactical situation displays, modular processing, tactical communications and defensive management decision aids in an F/A-18 environment as well as the integration of Phase I offboard components and initial distributed simulation internetting. This will be followed by an F/A-18 testbed flight demonstration of technologies proven in simulation.
- **Phase III** - Expansion of the Phase II laboratory simulation capabilities to evaluate additional onboard/offboard functionality, and the application of synthetic environments technology. Like Phase II, laboratory simulations will be followed by flight worthiness testbed demonstrations as well as follow-on integration of laboratory helmet mounted display (HMD) functionality.
- **Phase IV** - In-cockpit laboratory and flight demonstrations of HMD and synthetic environments technology as well as combined RTS/JPP demonstrations, joint service technology demonstrations and fleet transition planning.

Core technical issues being addressed in the long term risk reduction plan include:

- **System Integration** - Onboard/offboard functional partitioning, software/database commonality, interface connectivity and interoperability.

- **Pilot Vehicle Interface** - Situation awareness, situation data display, responsiveness, workload reduction, control interface methods and pilot intent/monitoring.

- **Mission Management and Replanning** - Aircrew alternative acceptance, replan robustness, flyability and responsiveness.

- **Data Management** - Capture, filtering, correlation, retention and dissemination.

- **Core Technologies** - Advanced processing, display, communications, mass data storage, and display technologies as previously noted.

- **Security/Safety** - Sensitive data accessibility, onboard equipment and flight safety.

The key to ATIMS development is maintaining synergism with related DoD programs and avoiding duplication of effort by incorporating applicable technologies within the ATIMS scope. Consequently, ATIMS has coordinated with key Navy research agencies (e.g., NRL, NR&D, NAWCWPNS), non-Navy research agencies (e.g., Air Force Wright Laboratory, Army Night Vision Laboratory) and the Advanced Research Projects Agency (ARPA).

RESULTS In 1993, the ATIMS program started taking a look at the requirements of information management. The initial concept demonstration for an information management capability was done at the USAF Wright Laboratory Advanced Fighter Technology Integration (AFTI) avionics simulator hotbench at the Lockheed Fort Worth Company, Texas. This testbed provided a starting point to satisfy lessons learned during Desert Storm about information capability shortcomings of today's aircraft and the impact of RTIC on mission execution. The demonstration showed rapid Tactical Aircraft Mission Planning System (TAMPS) preflight data transfer via 1553 bus downloading, limited threat avoidance inflight replanning, and post-flight TAMPS mission playback for evaluation of planned versus actual performance.

Phase II began in 1994 at the high fidelity F/A-18 simulation environment at Boeing Defense and Space Group's Integrated Technology Development Laboratory (ITDL). This demo showed that significant processing power, high resolution displays, and increased communications throughput were needed to provide RTIC/RTOC capabilities. Additionally, the concept of passing imagery through JTIDS was explored during this demo and an understanding was developed that commercial technology was adaptable to many of ATIMS' needs.

ATIMS had the opportunity in 1995 to become involved in two exercises. In Roving Sands, ATIMS participated via data link from the ITDL domed F/A-18 simulator as a simulated wingman to a live F-15. ATIMS had its initial flight later in 1995 as it was integrated into an AH-64 Apache helicopter as part of the Army National Guard's Deeplook '95 exercise. These exercises allowed ATIMS to further refine the needs of the airborne piece of information management and gave an initial understanding of the ground station and C4I architecture requirements. Pilot feedback became very useful as well in properly defining the information needs in the cockpit.

ATIMS again participated in the Deeplook exercise in 1996 with two airborne systems in AH-64 Apaches, an ATIMS equipped track vehicle and an ATIMS ground station. This provided an arena to further refine the C4I issues associated with real-time information flow with multiple ATIMS equipped units and allowed an initial view of how the CONOPS for multiple systems should be formed. Additionally, in 1996 a development concept for integrating a "high end" ATIMS system into the F/A-18 was done by McDonnell Douglas Aerospace. This study was used to determine how to integrate ATIMS in the most economical way with the least impact to other aircraft systems. The "high end" ATIMS system design that resulted from this study is the Mission Management Technology (MMT) Insertion Device, which incorporates the use of advanced parallel processing and mission management algorithms through a VME design. This design uses on and off board information via data links and passive monitoring of the 1553 avionics bus (Figure 6). As part of the open architecture approach, the MMT was designed in a way that isolates it from flight control systems, providing the capability to upgrade the ATIMS hardware and software without incurring the extra costs associated with re-flight certifying the aircraft.

This year ATIMS has been building the MMT with the plan originally to put it in an AV-8B Harrier and now with plans to integrate it into an F/A-18C at NAS Fallon. In March, ATIMS participated in the Marine Corps Hunter Warrior exercise. For this exercise, ATIMS was installed in an AH-1W Cobra helicopter and an ATIMS ground station was integrated into the Marine Corps' Experimental Command Operations Center (ECOC). The ATIMS equipped Cobra met objectives as a sensor and a shooter on the battlefield. Through interfaces with onboard sensors, ATIMS showed the ability to capture an image of the enemy and send it along with coordinates to the ECOC in real-time. This information was easily reviewed within ECOC, providing a rapid means by which to verify targets of opportunity and give clearance to attack. Additionally, this year the initial demonstration of the synthetic environments technology through the incorporation of a helmet mounted display was demonstrated at the Boeing Defense and Space Group's F/A-18 domed simulator. This demonstration showed that the ATIMS avionics would support the helmet mounted display and that the technology was much more advanced than previously thought. The synthetic environments capability presented information to the aircrew in a way that was easily assimilated and provided the pilot a much greater "head out of the cockpit" capability, increasing situational awareness dramatically.

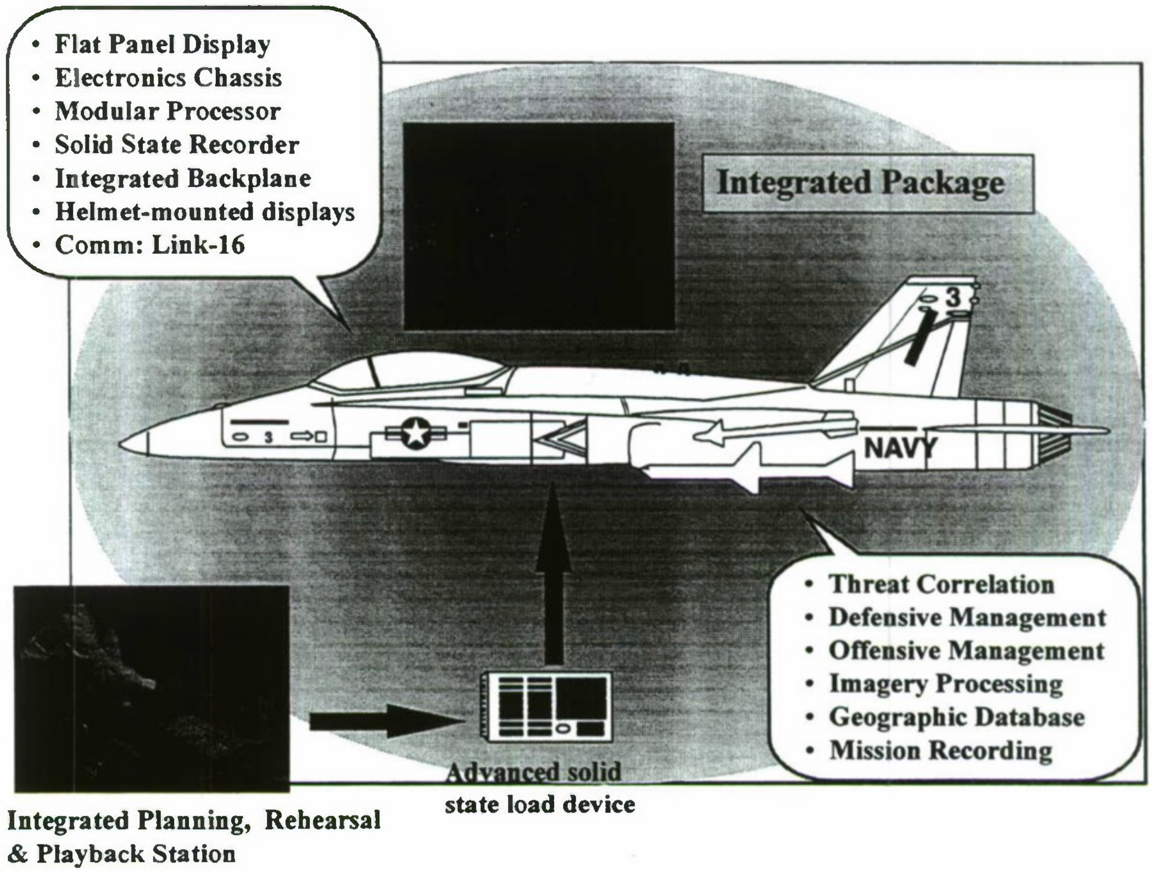


Figure 6. ATIMS fuses onboard data, offboard data, and mission planning databases with advanced processing and software to provide an integrated mission management capability.

The ATIMS program has successfully demonstrated many of the goals set forth in the four phased development approach. Many important lessons have been learned about information management along the way. In order to provide enhanced situational awareness, real-time mission management/execution, RTIC/RTOC, and other desirable operational improvements without overloading the pilot some of the lessons learned include:

- Processing power required is extremely high. The approximate requirements of the ATIMS MMT using a VME-64 design are 350-360 Million Instructions Per Second (MIPS) before additional processing requirements with the incorporation of a Helmet Mounted Display and Programmable Digital Radio.

- Parallel Processing instead of central processing to prevent latency of information.

- Distribute processing to reduce bus loading.

- Augment cockpit resources, i.e., advanced processors and algorithms can be used to assist pilot through kneeboard calculations and decision support.

- Real-time command and control participation in mission execution needed.

- SATCOM assures connectivity in all flight/terrain profiles.

- Existing aircraft architectures are problematic.
    - Legacy systems are too expensive to upgrade.
    - Existing bus structures and central computers are insufficient to support system throughput.

SUMMARY The ATIMS program is demonstrating the utility of advanced technologies to support the access of in-flight mission information updates for real-time mission replanning, improve pilot situational awareness and allow pilots to conduct strike rehearsal (Figure 7). The four-phased ATIMS development plan is conducting incremental laboratory and flight demonstrations which tie together applicable technologies in a system context and will enable the parallel spin-off of fleet insertion upgrade programs. In the coming years, the ATIMS program plans to flight test the MMT on a Navy platform that provides complex mission needs to properly define the requirements of information management in the aircraft and to develop the command and control information flow capabilities outside the aircraft. Additional work will be done to refine and develop the helmet mounted display technology and incorporate a Programmable Digital Radio (PDR) into the MMT. The PDR will provide connectivity to the crucial communications links of the "Digital Battlefield" (i.e., Link 16, Tactical Information Broadcast System (TIBS), Global Broadcast System (GBS), and UHF Data). The encompassing goal is to provide an integrated information warfare capability that will allow the exchange of real-time information between aircraft, ground vehicles, ground troops, and command and control of a multi-service and multi-national battlefield. This capability will enhance survivability, operational flexibility and mission effectiveness of the tactical aviator and enhance warfighting ability through out the battlefield.

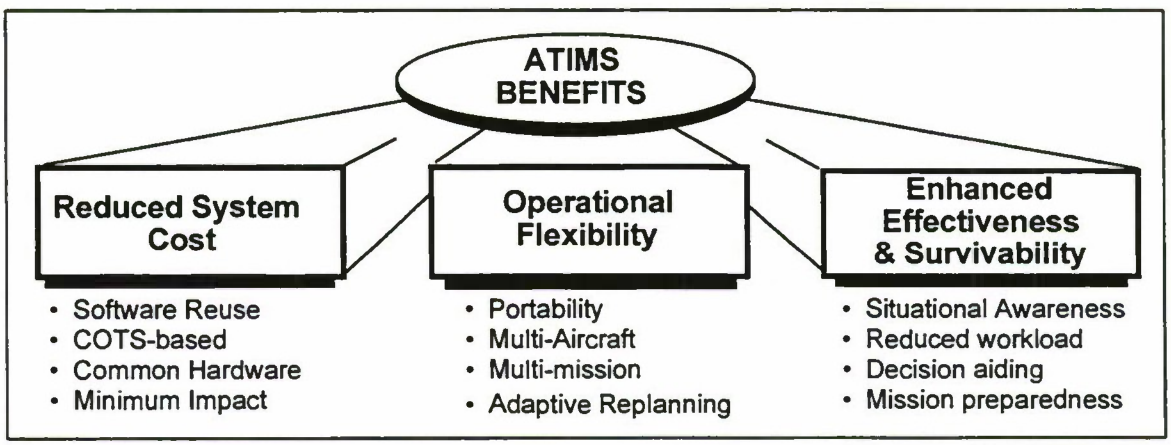


Figure 7. The ATIMS program is demonstrating candidate advanced technology for fleet insertion to simultaneously reduce avionics costs and increase mission effectiveness.

PHYSICS-BASED FEATURE DISCOVERY FOR TARGET RECOGNITION

Jonathan D. Michel
AbTech Corporation
Charlottesville, VA

Greg Arnold
Wright Laboratories AACA
Wright-Patterson AFB, OH

Abstract

Recent work in the area of Thermophysically Invariant (TPI) features of long wave infrared (LWIR) imagery led to the development of analytically derived model-based features that remain constant under a range of conditions. Limited practical success has been made in application and work is continuing in the development of the theoretical basis for a purely analytic approach. This paper will define an approach that makes use of the established thermophysical physics-based model that incorporates the thermal energy exchange of target features. Combining the physics-based approach with an empirical data modeling approach will leverage the potential for features that can be tested in ground-truthed experimentation or simulation. Adoption of an empirical approach, the use of non-parametric data models, will overcome specific drawbacks to the existing methods for analytic feature identification.

Introduction

This paper defines an approach which combines the physics-based analytical paradigm with a proven Automatic Target Recognition (ATR) data modeling methodology. The physics-based approach relies on the application of the conservation of energy relationship to the surface of the target and creates features dependent on the material composition of the target. Data modeling will add the advantage of an empirical examination of the feature behavior while considering features framed within an energy exchange model. Thus, the examination of feature behavior from exemplary images of the target will be anchored by the physics-based model of the target. The objective is to use the analytic model to gain the direct benefits of the physics-based approach while incorporating the advantages of data modeling to derive useful features.

The ATR problem considered in this work can be applied to several recognition problems. The problem considered here is site change detection. Site change applications are important in commercial and defense scenarios. A previous approach has attempted to analytically specify features of an energy exchange model. Invariant image features were identified by deriving features not dependent on the driving conditions of the imaging scenario. However, these analytic models are often coarse approximations of the real physical phenomenon. The exact physics of the scene are unknown and the assumptions necessary to interpret the scene and form features contribute to error in this implementation of a purely analytic approach. Thus, the features derived have systematic errors skewing the results and contributing to lower accuracy in the recognition process. In this work we apply a learning algorithm paradigm to learn from examples of the feature formation process. By applying supervised learning techniques the basic physical phenomenon can be captured in features while compensating for the approximate physical model error.

Statistical Networks are a natural choice for the selection of a learning algorithm to form features capable of distinguishing change in scenes. Their ability to form arbitrarily complex decision surfaces within a feature space will contribute to finding optimal relationships. This project will involve exploring the large space of possible non-parametric polynomial network models.

The approach described here is based on recent research that has defined methods for creating image-derived LWIR object features that vary only when the parameters of the object composition materials change [1]. Employing a hypothesize and verify paradigm, these features are based on interpretation of the conservation of energy equation. One such interpretation creates a linear form, which is then assumed to undergo a constrained linear transformation of the *known* coefficients from scene to scene. Using this system, several invariant features can be formed that are expected to be stable from scene to scene and distinct when a different class of objects are imaged. In tests on real data, the results for these features were ambiguous. For example, in one test on real data, 28 of the 58 predicted invariant features had a

quality ratio (std/mean) greater than 1 (poor stability), 21 greater than 2 (worse stability), and 12 greater than 3 (very poor). Of the 30 features computed to be invariant, only a few were reasonably stable and gave good interclass separability – the feature applied to another target class gave distinct results. Specifically, 30 features had a quality ratio less than 1 (reasonable stable). But only three of the features had good stability **and** were distinct when applied to other classes.

The ambiguous results from the analytically derived forms can be a product of two factors. (1) The assumptions invoked by the physical heat flow model can be a source of error (e.g., emissivity is constant, lateral conduction is negligible, the adjacent material to the spot imaged has equivalent material parameters, etc.); and (2) The assumptions invoked by the transformation model can be a source of error. The fact that the independence of the variables is not assured may render the linear transformation assumption inappropriate.

There are two paths to take from this juncture; one is to further explore the forms of the conservation equation and apply invariance theory existing or new. *An alternate approach is to apply empirical methods to the problem using machine learning techniques.* Both approaches are valid and desirable. One will explore the forms of the conservation equation and apply invariance theory existing or new. *The alternate approach, presented here, is to apply empirical methods to the problem using machine learning techniques.* Both approaches are valid and desirable.

Background - A Thermophysical Approach to IR Image Analysis

An intuitive approach to thermo-physical interpretation of LWIR imagery is given in [2]. This approach rests upon the following observation, termed the "Thermal History Consistency Constraint" and analogous to Lowe's well known Viewpoint Consistency Constraint [3]: "The temperature of all target features for a passive target must be consistent with the heat flux transfer resulting from exposure to the same thermal history." In [2], this constraint is exploited by analyzing objects to locate components that are similar in terms of thermo-physical properties and then examining a temporal sequence of calibrated LWIR data to experimentally assess the degree to which such thermo-physically similar components exhibit similar temperature state temporal behavior. Such analysis was shown to lead to formulation of simple intensity ratio features exhibiting a strong degree of temporal stability that could be exploited provided:
(1) thermally homogeneous regions in the LWIR image corresponding to the thermo-physically similar object components could be reliably segmented, and (2) a target-specific geometric reference frame is available in order to correctly associate extracted regions with candidate object components.

To avoid the difficulties inherent in assumptions (1) and (2) above, an alternative technique applicable to overall object signatures was suggested in [2]. An analysis of typical LWIR `lumped parameter' object temperature modeling approaches suggests that over small time scales object temperature can be crudely modeled by a small dimensional linear system with algebraically separable spatial and temporal components. Ratios of spatial integrals of temperature with a simple set of orthonormal 2D polynomials (obtained from applying Gramm-Schmidt to 1, x, y, xy, x^2y, xy^2, x^3 and y^3) were tried. Some of the resulting functions were nearly constant with time when measured against 24 hours of LWIR imagery of a complex object (a tank) but no experimentation was done with multiple objects to examine between and within-class separation, so little can be drawn in the way of a substantive conclusion with respect to utility as an object identification technique.

<u>Thermophysical Model</u> (from [3,4]) – Consider an infinitesimal volume at the surface of the image. Energy absorbed by the surface equals the energy lost to the environment. The objects considered in this work are passive, i.e., there are no internal heat sources.

$$W_{abs} = W_{lost}$$

Energy absorbed by the surface (per unit surface area) is given by:

$$W_{abs} = \alpha_s W_i \cos(\Theta_i),$$

where W_i is the incident solar irradiation on a horizontal surface per unit area and is given by available empirical models (based on time, date and latitude of the scene) or by measurement with a pyranometer; Θ_i is the angle between the direction of irradiation and the surface normal; and α_s is the surface absorptivity which is related to the visual reflectance ρ_s by $\alpha_s = 1 - \rho_s$. Note that it is reasonable to use the visual reflectance to estimate the energy absorbed by the surface since approximately 90% of the energy in solar irradiation lies in the visible wavelengths [6]. The energy lost by the surface to the environment was given by:

$$W_{lost} = W_{cnd} + W_{st} + W_{cv} + W_{rad}$$

where W_{cnd} denotes the energy (per unit surface area) convected from the surface to the air which has temperature T_s and velocity V; W_{rad} is the energy (per unit surface area) lost by the surface to the environment via radiation; and W_{cnd} denotes the energy (per unit surface area) conducted from the surface into the interior of the object. The radiation energy loss is computed from:

$$W_{rad}^{-} = \varepsilon\sigma(T_s^4 - T_{amb}^4)$$

where σ denotes the Stefan-Boltzmann constant; T_s is the surface temperature of the imaged object; and T_{amb} is the ambient temperature. Assume ε for the atmosphere is equal to ε for the imaged object. This assumption is reasonable if the objects are not uncoated metals. This assumption may not hold if the imaged surface is exposed or an unoxidized metal which is usually rare.

The convected energy transfer is given by

$$W_{cv} = h(T_s - T_{amb})$$

where h is the average convected heat transfer coefficient for the imaged surface, which depends on the wind speed, thermophysical properties of the air, and surface geometry [6].

Lateral conduction from the elemental volume at the surface is assumed negligible since the temperature of the material adjacent to the surface volume under consideration may be assumed to be similar. In general, the internal temperature of the material will be different from that at the surface. The energy flow, due to this gradient, is expressed as the conducted energy:

$$W_{cnd} = k\, {}^{dT_s}\!/_{dx}$$

where k is the thermal conductivity of the material and x is distance below the surface.

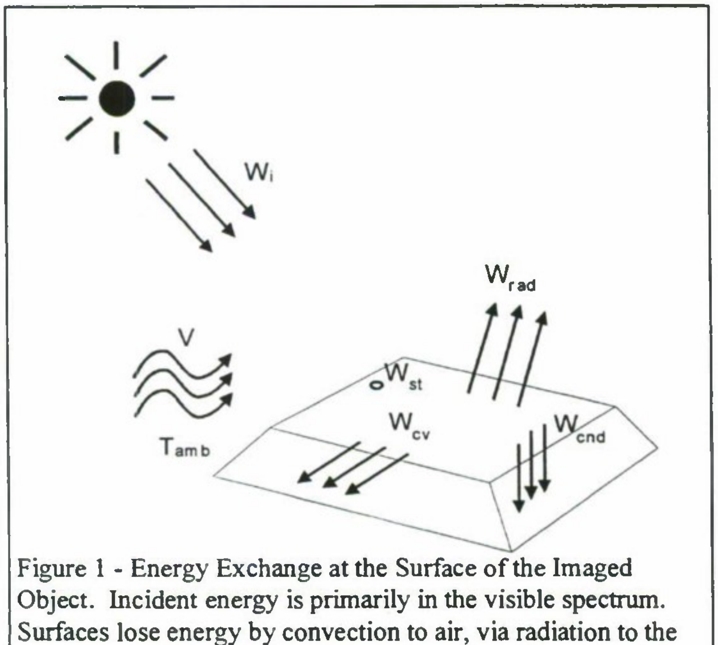


Figure 1 - Energy Exchange at the Surface of the Imaged Object. Incident energy is primarily in the visible spectrum. Surfaces lose energy by convection to air, via radiation to the atmosphere, and via conduction to the interior of the object.

Within the elemental volume, the temperature is assumed uniform, and the increase in the stored energy given by:

$$W_{st} = C_T\, {}^{dT_s}\!/_{dt}$$

where C_T is the thermal capacitance for the material of the elemental surface volume. This is given by C_T = DVc, where D is the density of the surface material; V is the elemental volume; and c is the specific heat. Again, W_{st} is expressed in units of energy per unit surface area.

The approach described in [7] builds upon the approach defined in [5] where recognition features are derived from algebraic invariance theory. The approach presented in this section uses the form of the energy balance equation established above. The method of algebraic elimination is used to derive invariant features [8]. Examination of the elements

of the measurement vectors and the physical process by which the elements change from scene to scene produces the insight that the transformation is limited to a subgroup of *GL(5)*. Thus, invariants are available using fewer points than the approach described in [6]. An overview of the general algebraic elimination method for finding invariant functions of polynomials is given in [8]. The work in this section has been reported in [4,5,7,8].

The energy balance equation,

$$W_{lost} = W_{cnd} + W_{st} + W_{cv} + W_{rad}$$

is written in the following linear form:

$$a_1x_1 + a_2x_2 + a_3x_3 + a_4x_4 + a_5x_5 = \vec{a}^T\vec{x}$$

Using the expressions for the various energy components as presented above we can express each term in the above expression as:

$$\begin{array}{ll} a_1 = C_T & x_1 = -\,d_{Ts}/dt \\ a_2 = k & x_2 = -d_{Ts}/dx \\ a_3 = -T_s - T_{amb} & x_3 = h \\ a_4 = -\,\sigma T_s^4 - T_{amb}^4 & x_4 = \varepsilon \\ a_5 = \cos(O_T) & x_5 = W_I\,\alpha_s \end{array} \qquad (1)$$

Note that a calibrated LWIR image provides radiometric temperature. However, this requires knowledge of the emissivity, ε, of the surface. For common outdoor materials, and common paints and surface coatings, the value of ε is around 0.9 [9], [10]. Therefore, the radiometric temperature, T_s, may be computed based on the assumption, $\varepsilon = 0.9$. Hence, a_3 and a_4 can be computed from the LWIR image alone (and knowledge of the ambient temperature), while a_1, a_2 and a_5 are known when the identity and pose of the object is hypothesized. The "driving conditions," or unknown scene parameters that can change from scene to scene are given by the x_i, $i = 1, \ldots, 5$. Thus, each pixel in the thermal image defines a point in 5-D thermophysical space.

Consider two different LWIR images of a scene obtained under different scene conditions and from different viewpoints. For a given object, N points are selected such that (a) the points are visible in both views, and (b) each point lies on a different component of the object which differs in material composition and/or surface orientation. Assume (for the nonce) that the object pose for each view, and point correspondence between the two views are available (or hypothesized). A point in each view yields a measurement vector $a = \{a_1, a_2, a_3, a_4, a_5\}^T$ with components defined by eqn 1 and a corresponding driving conditions vector $\mathbf{x} = \{x_1, x_2, x_3, x_4, x_5\}^T$. Let a collection of N of these vectors compose a $(5XN)$ matrix, $A = \{\mathbf{a}_1, \mathbf{a}_2, \ldots, \mathbf{a}_N\}$ for the first scene/image. These same points in the second scene will define vectors that compose a $(5XN)$ matrix, $A' = \{a'_1, a'_2, \ldots, a'_N\}$. The driving condition matrix, $X = \{\mathbf{x}_1, \mathbf{x}_2, \ldots, \mathbf{x}_N\}^T$ from the first scene and $X' = \{\mathbf{x}'_1, \mathbf{x}'_2, \ldots, \mathbf{x}'_N\}^T$ from the second scene, are each of size $(NX5)$.

Here is where this new approach diverges from the previous work. Our past approach used the features described above of the LWIR imaged materials by considering the form of the transformation from one scene to the next and then deriving features that are invariant to that transformation. Rather than pursuing features by analytic consideration of the invariance to the transformation, we will consider forming primitive features in thermophysical feature space then empirically searching the space of polynomials of these primitives to find recognition features according to the desired criterion (e.g., low variance, strong inter-class separability, tolerance to noise, etc). In this approach we include the features considered in the past work since we are evaluating a much larger set of recognition features, including those found by the analytic methods. The past work led us to the development of features that were purely ratios of determinants. This work will use the determinant as a basis for features as the primitives; however, we will explore a larger class of functions to form the recognition features. Indeed, it is well known that invariants of linear forms are always functions of the powers of determinants [11].

The derivation of the features and the approach used here in the recognition system, as described above, relies on a number of assumptions. These assumptions are summarized. (1) Five points are chosen such that measurement vectors (in each scene) are linearly independent, i.e., one (or more) of the four points has different material properties and/or surface normal. (2) The points are corresponded. (3) Object identity is hypothesized (which is verified or refuted by the feature value). (4) The thermal capacitance and conductance of the object (surface) do not change while other variables are allowed to vary from one scene to another. (5) Emissivity of the imaged surface in the 8μm - 12 μm band is 0.9. The first assumption is met when the vectors are linearly independent by proper choice of the points, i.e., model formation. Most objects and scenes have sufficient diversity in surface material types to easily satisfy this requirement. The point correspondence and the hypothesis assumptions are satisfied in the method of application of the features. Assumption (5) is satisfied except for bare metal surfaces and esoteric low emissivity coatings.

In order for the recognition features to be useful, not only must the values of the feature be stable to scene conditions, but the value must be different if the measurement vector is obtained from a scene that does not contain the hypothesized scene and/or if the hypothesized pose is incorrect. Since the formulation above takes into account only feature invariance but not separability, a search for the best set of points that both identifies the object and separates the classes must be conducted over a given set of points identified on the object. The search may be conducted over all the combinations of the points in a set or until an acceptable feature is found. We have examined all combinations, first rating each set for their intra-class invariance, then further evaluating it for inter-class separability. Results on real LWIR data are described in the last section.

Approach

To reiterate the problem, *what LWIR data derived function yields target interclass separability and intraclass stability over the given range of operating conditions?* For a LWIR spectrum sensor, the primary contribution to gray level value is emitted energy through radiation exchange between the IR camera and the target surface. Thus, the gray level corresponding to an imaged surface has a dependence on the target surface temperature and is effected by the material composition, emissivity, and thermal history.

In a hypothesis and verify algorithm, as described here, the empirical approach to modeling presents the problem that if the hypothesized values are used in training the model, the recognition system will be skewed by the "constant" features. In essence, the algorithm will always answer that your hypothesis is correct. It will (properly) ignore the measured values. The approach used here to overcome this dilemma of incorporating a hypothesized physical model is to consider the determinant formed by following a set of points formed by both hypothesized and measured values of the balance of energy equation. The components of the relationship, as described above, are the heat flux terms relating the energy exchange at a point on the surface of the object. The components are listed in two sets:

$$(1) \qquad s_1 = \{C_T, k, T_s, T_{amb}, \cos(\Theta), \sigma\}$$

$$(2) \qquad s_2 = \{\alpha, \varepsilon, W_i, {}^{dT_s}\!/_{dt}, {}^{dT_s}\!/_{dx}\}$$

The set, s_1, contains data and parameters known from image measurements or from the hypothesis of the class of the vehicle. The set, s_2, contains *driving conditions* of the imaging process. A function composed of elements in set 1 from one or more points on the object will be computable from a single image (and model-base knowledge). The question: what function meets the criteria of a useful object recognition feature – the (often) competing criteria of strong intraclass stability and high interclass separability. This answer will be solved by modeling the class membership functions of thermophysical features using Statistical Networks. The following section gives a brief introduction to the modeling technology used in this work.

Statistical Networks™ Technology – Supervised machine learning approaches have shown encouraging preliminary success for many modeling tasks. Here, relational models (those which relate a set of inputs or

observations to a desired parameter estimate) are learned inductively from empirical evidence. Usually, relationships which potentially represent a complex process or environment are hypothesized and "scored" according to some criteria which minimizes error. On the basis of the performance of the hypothesized relational model, several refinements and adjustments are made based on a learning mechanism. Traditional statistical regression and neural network approaches offer some utility, but suffer from practical limitations which have been well documented [12]. ModelQuest results from three decades of application research and development in several fields including statistical modeling, uncertainty management, genetic algorithms, and neural networks. It demonstrates superior performance for target identification, forecasting, space object detection and classification, signal analysis, sensor fusion, aircraft vulnerability analysis, decision aiding, on-board aircraft diagnostics, financial analysis, non-destructive inspection, and medical evaluation.

Statistical Networks process information with complex mathematical functions. Functions are attractive because they capture a large number of complex relationships in a very compact and rapidly executable form [13]. The statistical learning algorithm discussed below produces a network of functional nodes – each node containing a multiple-term polynomial relationship. Polynomial nodes are an extremely powerful method for performing complex reasoning tasks – they are the basis of traditional neural networks and other modeling techniques. They process one, two, or three inputs to compute an output value; and contain a bias or constant term (w_0), and linear, quadratic, cubic, and cross terms. A LINEAR node processes several inputs and contains only the linear and bias terms. The equations for each node type are:

$$\text{SINGLE} = w_0 + w_1x_1 + w_2x_1^2 + w_3x_1^3$$
$$\text{DOUBLE} = w_0 + w_1x_1 + w_2x_1^2 + w_3x_1^3 + w_4x_2 + w_5x_2^2 + w_6x_2^3 + w_7x_1x_2$$
$$\text{TRIPLE} = w_0 + w_1x_1 + w_2x_2 + w_3x_3 + w_4x_1^2 + w_5x_2^2 + w_6x_3^2 + w_7x_1^3 + w_8x_2^3 + w_9x_3^3$$
$$+ w_{10}x_1x_2 + w_{11}x_1x_3 + w_{12}x_2x_3 + w_{13}x_1x_2x_3$$
$$\text{LINEAR} = w_0 + w_1x_1 + w_2x_2 + \ldots + w_nx_n$$

An example Statistical Network is shown in Figure 2. It is a feed-forward network of polynomial nodes processing information from left to right. Each node produces intermediate information which is used as inputs for subsequent nodes. This networking strategy segments the overall relationship being modeled into more manageable components, and simplifies the learning process. Functional networks are synthesized automatically from a flat file database where each column is an input or output parameter (e.g., a variable), and each row contains an example set of the parameters. A hypothesize and test strategy finds the network which best represents the relationships contained in the database.

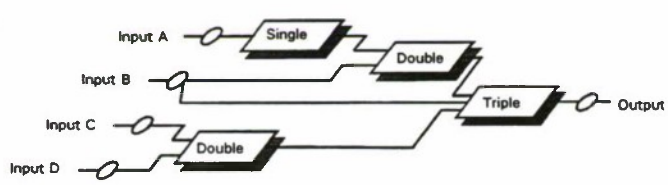


Figure 2 – Example Statistical Network

While individual nodes only allow up to three inputs and are limited to third order terms, employing them in the networking strategy shown in Figure 3 allows the overall network to accept any number of inputs. In addition, because a specific node can contain a third order term, a two-layer network can model a ninth-order relationship. An additional layer allows the modeling of up to 27th order relationships, etc. Therefore, networking relatively simple node types creates a powerful knowledge representation [14].

The ModelQuest learning process produces networks of functional elements which more effectively "learn" complex relationships among features than is often practical with other methods. This modeling paradigm

is described in detail in Reference [14]. The key to any machine learning strategy is the learning algorithm itself. It must be able to ***generalize*** from, and not ***memorize***, numerical examples of a problem domain. It must be able to automatically discover relationships to produce a model which performs well for not only training data but independent (i.e., real-world) data. The driving reason for this crucial requirement is ***that all data contain uncertainty***. Noisy, missing, conflicting, and erroneous data are all manifestations of uncertainty in numerical examples.

An effective machine learning algorithm must ***learn relationships*** and avoid ***memorizing noise*** in an ***automated*** manner. Statistical Networks achieve this through the use of intelligent search heuristics to find the optimal network architecture and a modeling criterion to ensure generalization. Following is a top-level summary of the ModelQuest learning algorithm (outlined in Figure 4):

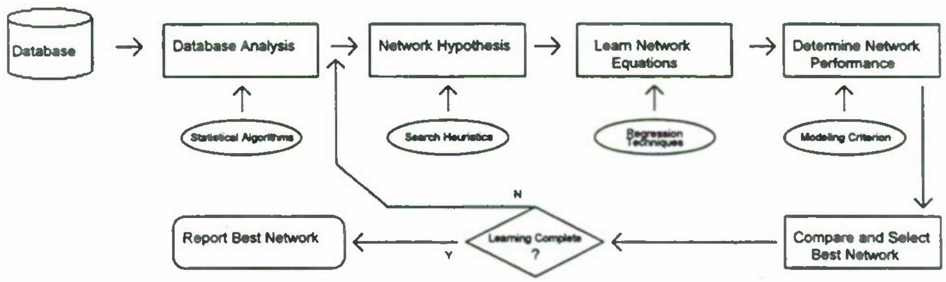


Figure 3 – Statistical Network Learning Algorithm

<u>Step 1</u>: Several statistical measures are computed for each database variable such as their mean and standard deviation. The values for each variable are normalized so that they exhibit a mean of zero and a standard deviation of unity, greatly enhancing node regression performed in Step 3.

<u>Step 2</u>: Candidate network architectures are hypothesized using graph-tree network search heuristics. The heuristics employ a survival of the fittest strategy – similar to the underlying concept of genetic algorithms – by hypothesizing more refined versions of networks that have already exhibited promise. Initially, very simple network models are hypothesized (i.e., those which contain only one node). The best of these simple models (as scored by the modeling criterion in Step 4) are then used along the original input parameters as building blocks to hypothesize more complex networks. Search heuristics determine the best manner to combine simpler networks to form more complex ones. This process is repeated (automatically) several times, each providing an additional network layer.

<u>Step 3</u>: For each hypothesized network, coefficients for each node are determined using advanced regression algorithms. The result of Step 3 is values for the coefficients ($w_o, w_1, w_2, \ldots, w_n$) in each network node.

<u>Step 4</u>: Each network is "scored" with the Predicted Squared Error (PSE) modeling criterion, shown in Figure 5. The PSE was developed at Stanford University in the early 1980's specifically as modeling criterion for statistical learning [17]. The network with the best (i.e., least) score is selected as the best for a particular database. The PSE performs a trade-off between network complexity and accuracy to find the simplest network that best models training *and* independent data. It gives an analytic estimate of the network for independent data. The PSE is:

$$PSE = FSE + KP = FSE + CPM\,[\,(2K/N) * s_p^2\,], \text{ where,}$$

- *FSE* is the fitting squared error of the network on the training data.

- *KP* penalizes more complex networks, as they are more likely to overfit training data and, therefore, not perform well on independent data.
- *CPM* is a complexity penalty multiplier, used to vary the emphasis of the *KP* term.
- *K* is the total number of network coefficients in the network model.
- *N* is the number of training observations.
- s_p^2 is an a priori estimate of the optimal model's error variance.

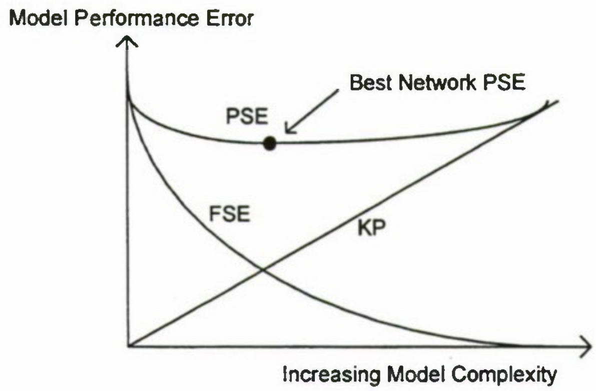


Figure 4 – PSE Modeling Criterion

The PSE produces networks which avoid modeling noise and overfitting training data. The Statistical Network synthesis process begins at the left of the PSE curve shown in Figure 5. As the complexity of hypothesized networks increases, the PSE of those networks decrease until the network with the minimum PSE is found. The learning process ends when certain "stopping criteria" are met (Figure 4). These criteria include heuristics which recognize that the learning process is taking place on the upward slope of the PSE curve, and that the best network has been found.

While the statistical network is parametric at the node level, the hypothesis heuristics and modeling criterion at the network level create an *automated* non-parametric process. Therefore, the human user is *not* required to be an integral part of the learning algorithm as *is* required by other approaches. This allows the system developer to focus limited resources on other issues, such as problem definition, system design, model evaluation, and system integration.

Method - Thermophysical Modeling

A model-based hypothesize and verify algorithm as described in [1] asserts material properties for a given imaged point on the object. For a given point, a feature vectors is formed by combining the hypothesized data with the measured data types, e.g.,

$$\mathrm{v} = \{C_T, k, T_s, T_{amb}, \cos(\Theta)\}$$

where the material properties are hypothesized and the measured quantities are derived from the image or the ambient conditions.

The approach for creating model-based features in a hypothesize and verify algorithm is to use the hypothesized values as a reference frame for the measured values, then compute features based on the measured values within the hypothesized values frame. For example, consider using three components, $\mathrm{v}' = \{C_T, k, T_s\}$. The measured value, surface temperature can be cast in the frame created by the material properties: thermal conductance and thermal capacitance. Using three points the components form a parallelopiped in the three dimensional space, the volume of which changes as the surface temperature of the points change, the other two components are considered constant. Given N points on the object and k components of the thermophysical vector, there are $\binom{N}{k}$ features available.

An example of the advantage of our nonparametric modeling method is that highly correlated and dependent features can be used as inputs without the reservation of skewing the model as would be the case in a standard modeling scheme. Thus, we can span a larger set of potential features. *Note: A generalized form of the polynomial features found analytically in [1] will be within the search space of the modeling*

algorithm – these features will be processed and compared statistically to other features. The generalization is that each term will have a coefficient weight found in the statistical network search.

It is evident that given the large number of features and a limited set of example imagery, it is critical to reduce the dimensionality of the input space for target modeling to avoid the creation of a highly biased model. The data modeling algorithm described in above section will allow the exploration of a huge number of potential models and feature combinations. Finding accurate models with good predictive ability will be possible. The highly automated model search technology will allow the experimentation with many different physics-based models.

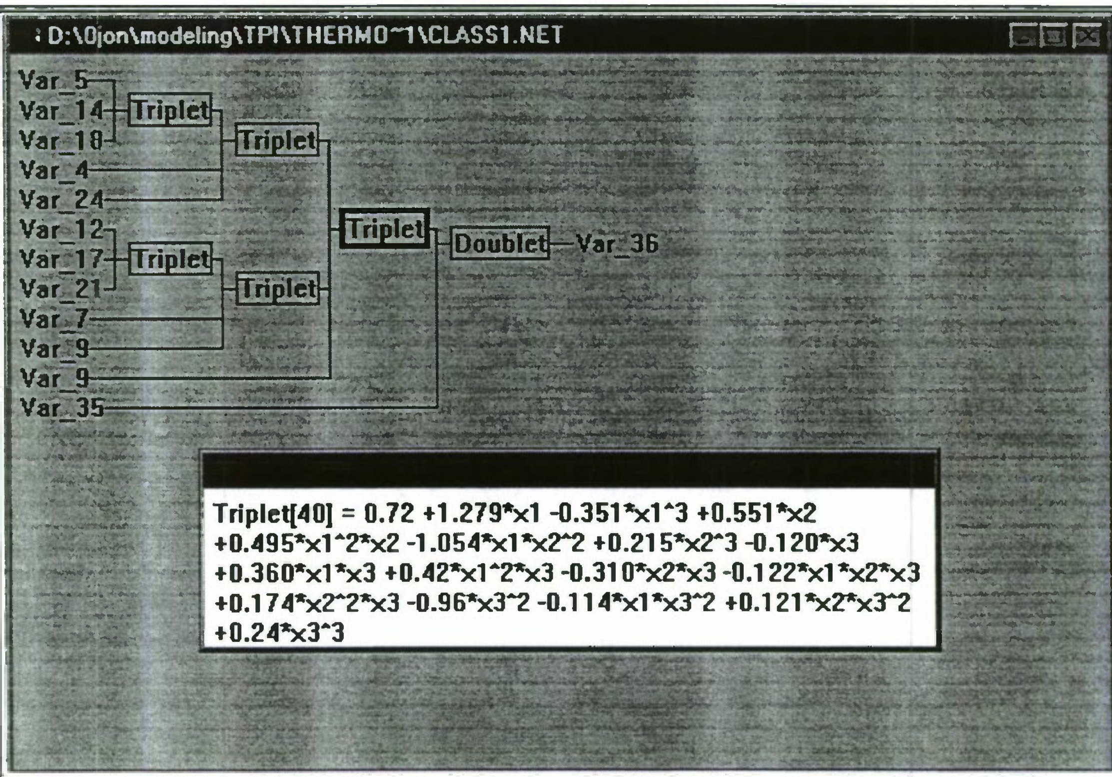


Figure 5 - ModelQuest Polynomial Network found empirically from data over a four day period of observing a scene. Seven material types were used to characterize the scene. Each material contributed a four dimensional point in the thermophysical space. The (7 choose 4) combinations of determinant values formed by the four dimensional vectors are the 35 variables used to form the feature. Note that ModelQuest algorithm automatically balanced the bias and variance of the model by determining the best network for the problem., eliminating variables and constructing the network that best balances fitting square error and model complexity. The inset shows the polynomial equation that combines the variables in the third layer.

Results:

A feature computation that is suitable for use in an object recognition system that employs a hypothesize-and-verify strategy is described here. The scheme would consist of the following steps:

- Extract geometric features in the LWIR image, e.g., lines and conics.
- For image region, *r*, hypothesize scene class, *k*, and pose using, for example, geometric invariants as proposed by Forsyth, et al. [16].
- Use the model of object *k* and project visible points labeled i =1,2,...onto image region *r* using scaled orthographic projection.

- For point labeled i in the image region, assign thermophysical properties of point labeled i in the model of object k.
- Use the gray levels at each point and the assigned thermophysical properties, to compute the measurement matrices A and A', and hence compute the feature, $f^k(r) = F^K(A_1, A_2, \ldots, A_N)$.
- Compare feature $f^k(r)$ with model prototype F^K to verify the hypothesis.

In one example of such a task, images are obtained, periodically, by an aircraft or satellite flying over a site to be monitored. The imagery is compared with known prior information, or detailed site models, to determine if any changes have occurred. For example, it may be important to detect if a patch of gravel, or dirt, has been replaced with a concrete, or asphalt, surface at some factory or construction site being monitored. Since some information is usually available of the site being monitored (in the form of site models), and also of the imaging parameters of the sensors, the site change detection task follows the paradigms of context-based vision and model-based vision. It is also closely related to the task of object recognition where a hypothesis of an object composed of different material types is verified or refuted. A particular site may be considered to be a composition of a specific collection of materials. A feature may be used to verify the existence of this composition. A change in the site should result in a feature value other than that expected for the original site.

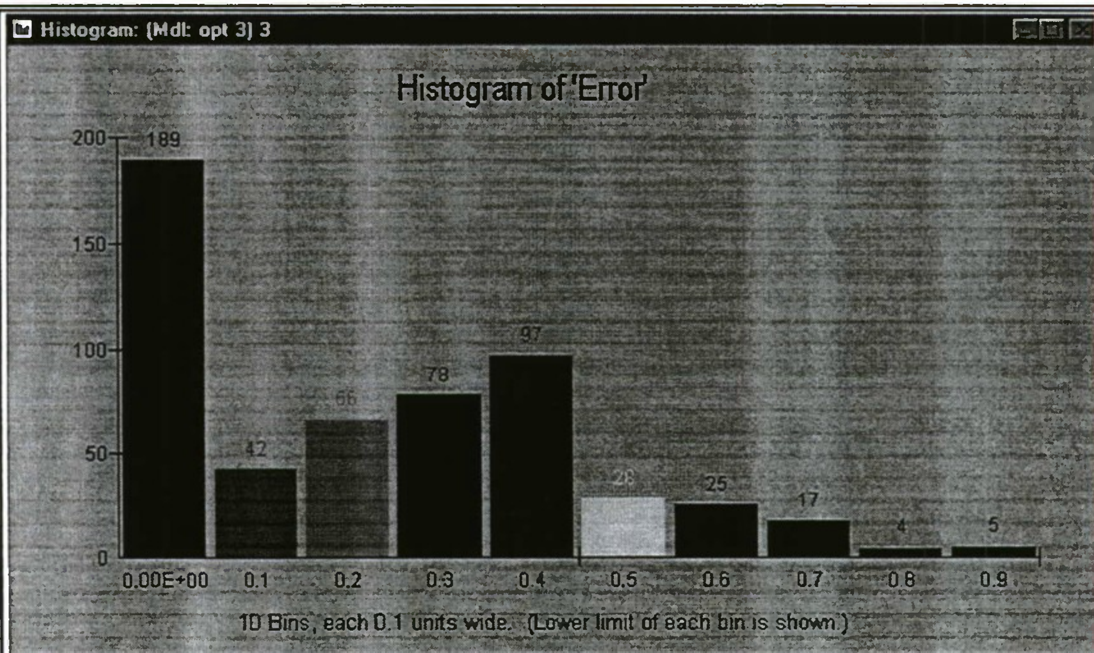


Figure 6 - The histogram shows results from scene change detection experiment using thermophysically-based features. The experiment held 540 randomly selected observations of the 3456 total as an evaluation set. Using a simple 0.5 threshold for the decision criteria, the error rate for site change detection in noisy data is 0.15.

The application of this scheme for site change detection is straightforward. For a site being monitored, M different types of surfaces are selected *a priori* to produce a stable thermophysical feature. Note that we must have $M > 4$. One may also be able to specify more than one feature, and, hence, establish a feature vector. An LWIR image is first registered with the site using established techniques [16]. The gray levels from the M selected regions along with the known material properties are used to generate the TP features. When one or more of the surfaces change (e.g., from gravel to concrete) then the feature (vector) computed from the scene under the hypothesis of the prior material types will produce a value different from that expected. The detection of the change is, thus, linked to the refutation of an incorrect hypothesis. The method of computing empirical thermophysical features discussed above was applied to real LWIR

imagery acquired at different times of the day. An outdoor scene was considered, several points were selected on the surfaces of different materials and/or orientation. The measurement tensor was computed for each point, for each image/scene.

Conclusion

This paper defines an approach which combines the physics-based analytical paradigm with a proven data modeling methodology. The physics-based approach relies on the application of the conservation of energy relationship to the surface of the target and brings to bear the advantages of a model-based approach. Approach the problem using data modeling methods add the advantage of an empirical examination of the feature behavior, considering features framed within energy exchange model. Thus, the examination of feature behavior from exemplary images of a target will be anchored by the physics-based model of the target.

The objective is to use the analytic model to gain benefits of the physics-based approach while incorporating the advantages of data modeling to derive useful answers. An analogy is searching for a goal in an unknown maze. The use of a coarse analytic model is like knowing the general direction of the goal. We can greatly increase the accuracy of the goal search by knowing the general direction. We can derive stable and useful features by framing the feature search within the approximate physical model.

References:

1. Michel, J., "Thermophysical Algebraic Invariance for Infrared Image Interpretation," *Ph.D. Dissertation*, Department of Electrical Engineering, University of Virginia, May 1996.
2. M.J. Gauder, V.J. Velton, L.A. Westerkamp, J. Mundy, and D. Forsyth, "Thermal Invariants for Infrared Target Recognition," *ATR Syst. Technol. Conference,* 1993.
3. D.G. Lowe, "The Viewpoint Consistency Constraint," *Int. J. Comput. Vis.*, Vol 1, No 1, pp 57-72, 1987.
4. Jonathan Michel, N.Nandhakumar, and V.Velten, "Thermophysical Algebraic Invariants from Infrared Imagery," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, Vol 19, No. 1, January 1997.
5. N.Nandhakumar, Jonathan Michel, Greg Arnold, G.Tsihrintzis, and V.Velten, "Robust Thermophysics-based Interpretation of Radiometrically Uncalibrated IR Images for ATR and Site Change Detection,"*IEEE Transactions on Image Processing*, Vol 6 No. 1, January 1997.
6. Incropera F.P., DeWitt, D.P. *Fundamentals of Heat Transfer,* John Wiley and Sons, New York, 1981.
7. Jonathan Michel, N. Nandhakumar, T. Saxena, and D. Kapur, "Using Elimination Methods to Compute Thermophysical Algberaic Invariants from Infrared Imagery," *The Thirteenth National Conference on Artificial Intelligence*, AAAI, Portland, Or. August 4-8, 1996.
8. Jonathan Michel, N. Nandhakumar, T. Saxena, and D. Kapur, "Geometric, Algebraic and Thermophysical Techniques for Object Recognition in IR Imagery," submitted to *CVIU.*
9. S.B. Idso and R.D. Jackson, "Thermal Radiation," *Geophysical Research,* 82(28), September 1977.
10. Y.S. Touloukian and C.Y. Ho, Eds, *Thermal Radiative Properties of Nonmetallic Solids,* Thermalphysical Properties of Matter, Vol. 8, Thermophysical Properties Resaerch Cneter of Purdue University, Data Series, Plenum Publishing Corp., N.Y. 1972.
11. G.B. Gurevich, "Foundations of the Theory of Algebraic Invariants," P. Noordhoff Ltd - Groningen, The Netherlands, 1964.
12. Montgomery, Gerard J., and Drake, Keith C., "Abductive Reasoning Networks," *Neurocomputing – An International Journal*, Volume 2, No. 3, June 1991, pages 97-104.
13. Farlow, S. J.(editor), *Self-Organizing Methods in Modeling: GMDH Type Algorithms*, Marcel-Dekker, Inc., New York, 1984.
14. Barron, A. "*Predicted Squared Error: A Criterion for Automatic Model Selection,*" Chap 4 of *Self-Organizing Methods in Modeling: GMDH Type Algorithms*, Farlow, S. J (editor) Marcel-Dekker, Inc., New York, 1984.
15. Michel, J., Drake, K., Kim, R. Engler, D., "ModelQuest Data Mining Toolbox in the Khoros Environment," *Proceedings of the Khoros Symposium,* Albuquerque, NM, April 1997.
16. Forsyth, D. Mundy, J. Zisserman, A. Ceolho, C. Heller, A. Rothwell, C., "Invariant Descriptors for 3D Object Recognition and Pose," *IEEE Transactions PAMI*, Vol 13, No 12, Oct 1991.

SITUATION AWARENESS USING AN INTEGRATED HELMET AUDIO-VISUAL SYSTEM

Robert K. Osgood
Naval Air Warfare Center
Weapons Division
China Lake, CA

Steve Adams
Naval Air Warfare Center
Weapons Division
China Lake, CA

The Joint Strike Fighter (JSF) Program Office successfully demonstrated technology that increased situation awareness (SA) for air-to-ground mission profiles. The Integrated Helmet Audio-Visual System (IHAVS) project was originated and funded by the JSF Flight Systems Integrated Product Team (IPT). IHAVS was conceived as a short-duration technology maturation and flight demonstration program. The synergistic effort and cooperation of joint government and industry organizations was a key factor in the success of the demonstration. The IHAVS project integrated head-mounted aural, verbal, and visual cockpit technologies, previously demonstrated individually, into a single system in a TAV-8B Harrier. A 25-event flight test operation was conducted by the Naval Air Warfare Center Weapons Division (NAWCWD) at China Lake, CA which demonstrated the utility of this integrated human systems interface for performing strike missions. IHAVS technologies included a General Electric Company (GEC) Viper II binocular helmet-mounted display (HMD) system, a Polhemus helmet tracker, a USAF Armstrong Laboratory 3-D audio system with active noise reduction (ANR), a Smiths Industries interactive voice module (IVM), and a Lockheed Martin Aeronutronics NITE Hawk Self-Cooled Forward Looking Infrared Targeting POD (TPOD). Aircraft integration was performed by McDonnell Douglas Aerospace and the NAWCWD, China Lake, CA. Three test pilots, one each from the Marine Corps, Air Force, and Navy, were selected to fly realistic air-to-ground missions including tactical ingress, threat management, and engagements against targets of opportunity. Objectives of the IHAVS program were to demonstrate a reduction in pilot workload while increasing survivability, lethality and SA. Results suggest that IHAVS reduced workload, increased pilot SA, and had the potential to increase survivability and lethality. IHAVS enabled the pilot to improve threat avoidance, conduct multiple sequential target acquisitions during attack, and simulate weapons delivery successfully.

(Reprint of executive summary; formal paper not available)

This page intentionally blank.

IMPACT OF C3 ON A/D FIGHTER EFFECTIVENESS

Approaches to Improve Situational Awareness
AND AAM Missile Employment

G. Wunderlich, IABG
Industrieanlagen-Betriebsgesellschaft mbH (IABG)
Einsteinstrasse 20
85521 Ottobrunn, Germany

ABSTRACT

Progress concerning increasing airborne weapon system performance requires extended aquisition and fire control capabilities to effectively use increased missile range (exceeding onboard sensor range) and potential increased sensor volume for improved situational awareness.

Three studies are presented dealing with these topics:

The first one was performed a couple of years ago as a joint SHAPE Technical Centre (STC) and IABG - effort to evaluate the impact of providing Air Situation Information via data link (MIDS/JTIDS) from an E-3A to air defence fighters. This enables the fighter pilot to increase situational awareness; to stay quiet, thus, not altering the target's warning receivers; and to execute an optimized approach towards the target, e.g., staying outside of the target's sensor volume. For this purpose, an STC-developed software module simulating E-3A detection and tracking was implemented into a German many-on-many air combat simulation model SILKA provided by IABG. The results were an increase of about 25% concerning enemies losses and a decrease of about 30% concerning own losses in the scenario chosen for the study.

The second study was performed recently by the 4 Power SNR Technical Group (TG) on Future Beyond Visual Range (BVR) AAM's (US, UK, FR, GE). This TG evaluates new technologies concerning airframe, propulsion, and seekers. Because some of the future AAM concepts exceed the onboard sensor range, the option of "Third Party Targeting" (TPT) was evaluated. For this pupose, the specifications concerning Track Accuracy Classes (TAC) for the NATO ACCS (ΔX, ΔY, ΔZ, refresh rate, track completeness, etc.) were implemented into a m vs n manned air combat simulation (CEV, France) as well as into m vs n unmanned air battle simulation (ARENA, UK). The study revealed large gains in using TPT, especially in case of adapting seeker look angle to the specified TPT accuracy. Further excursions and evaluations will be performed during the ongoing 4 Power work.

The third study was performed for a European customer and supports the first two studies.

PART 1: MIDS/JTIDS AIR COMBAT SIMULATION

1.1 Introduction

The quantification of the increases in effectiveness that might be provided by new weapons systems, avionics, etc., to a combat aircraft is a complex and sometimes controversial task. This is essentially due to the inevitable fact that any such analysis must be based on carefully chosen assumptions. Such considerations can be analysed and their effects determined with sufficient number of excursions and iterations. It is clear that, with the increasing cost and complexity of avionics, simulations will be a growing component of the decision making process. This presentation will present results of software simulations activity aimed at quantifying the effectiveness enhancement implications of an externally/remotely provided air situation picture, through an appropriate avionics-fit, to the pilot of a modern fighter/intercepter (Ref. 1). Results of other related simulation activities are available in Refs. 2-5.

MIDS/JTIDS (Multifunction/Joint Tactical Information Distribution System) is clearly such an avionics-fit and could be considered as the first example of the next generation of integrated CNI (Communication Navigation Identification) systems.

From the view point of the fighter/interceptor pilot (air defence missions are our primary concern), the implications of a MIDS like system could be stated as follows (Refs. 1,2).

- Obtain air situation around the pilot without having to turn radar on and perform surveillance manoeuvres.

- Increase surveillance volume.

- Receive, in real time, other classified information (data, link, voice).

- Display mission related information (navigation, tasking, status, threat, ...) visually ("parallel input") rather than communicate by voise ("serial input").

- Extend coverage under jamming due to the ECM resistance (particularly with respect to present UHF radios).

- Use relative position measurement capability to aid the identification of friendlies.

In terms of air defence operations, four generic aircraft control strategies or methods can be enumerated (in order of increasing information provided to the pilot):

a) Surveillance CAP (combat air patrol). No information provided to the pilot. Pilot acts according to information available from own sensors.

b) Broadcast or loose control. Some control information is available (e.g., raid size, direction ? ...). Pilot chooses actions based on this limited information.

c) Ground (or air) controlled intercept A/GGI, i.e., tight control. Pilot is controlled essentially close to the weapon release point.

d) MIDS/JTIDS: Pilot is provided the air situation from remote sensors and makes, within certain geographical or tactical boundaries provided by the control authorities, decision on how and when to approach combat.

Impact of C³ on Fighter Effectiveness

- Present status:

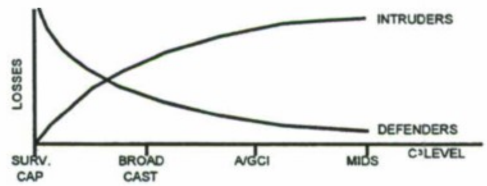
  depending on ECM/ECCM/ESM

  ★ A/GCI = best available option (voice controlled intercept)
  ★ Broadcast = option for low information level (intermittent voice information)
  ★ CAP = more or less no information (fighter has to rely on own sensors)

- Future possibility:

  use of a MIDS-like system will provide the fighter with more information than the best at present possible (A/GCI)

- Expected tendency:



- Evaluations performed:

  ★ CAP ➛ GCI
  ★ CAP ➛ GCI

Figure 1

The impact of C³ on fighter effectiveness is indicated in a more generic sense in Figure 1. A study performed for NATO AC-243 (Panel X) RSG-11 by STC/IABG, also using SILKA (Ref. 7), has indicated the significant effectiveness improvement implications of providing the pilot A/GGI information. This study compared essentially cases (a) and (c) above. Our study compares (c) and (d). The assumption of perfect UHF communications (currently the only means of A/GGI) is clearly an idealisation, as it is well known that the coverage available from the conventional UHF systems under ECM conditions will be very small and, hence, operationally extremely restrictive (Ref. 8). However, in order not to introduce another parameter (the level of UHF jamming) into the analysis it was found convenient to take the ideal benign case considered in the first generation ECCM UHF systems (e.g., HaveQuick, Colibri, etc.).

1.2 Software Air Combat Simulations

The SILKA software resident at IABG was augmented with STC developed subroutines to enable the simulation of exterally generated tactical air situation (TAP). Salient characteristics of SILKA "many on many" air combat software is given in Figure 2, and more information may be found in Ref. 6. As the external sensor providing the TAP to the fighters, the E-3A radar and tracking logic were modelled in a simplified but realistic manner, including range/azimuth/elevation errors and random omissions of plots. Figure 3 indicates the software utilised. The exact air situation data within SILKA is corrupted as shown before being passed onto the pilot decision logic. The provision of the TAP to the fighter/interceptor pilots was assumed to take place in real time as would be the case with MIDS/JTIDS (detection and tracking logic delays are of course simulated).

M/N-Combat Simulation

SILKA

CHARACTERISTICS/ASSUMPTIONS

- Max. 4AD-fighters, max. 26 opponents (FB, Escort)
- Same tactical principles for both sides
- Deterministic model
- Random processes for detection/missile effectiveness/multi-target generation
- Starting geometry, fuel, formation, etc., calculated in preliminary runs (A/GCI, CAP)
- Various A/C, missiles and avionics
- Statistical results
- Batch and MIL - version

Figure 2

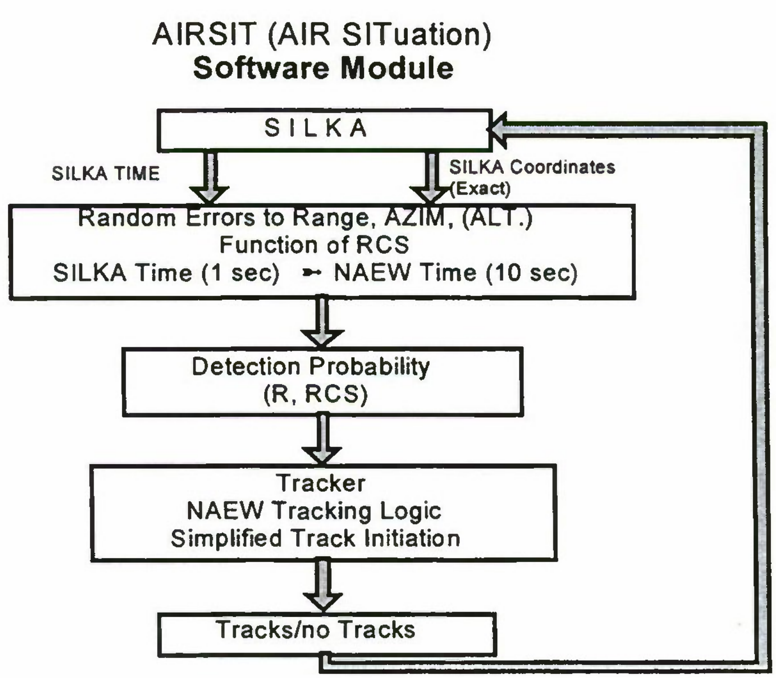


Figure 3

The TAP provided from external source through a protected (ECM and encryption) link, i.e., MIDS/JTIDS, has the following potential advantages (+) and limitations (-) (Figures 4a and 4b):

Potential Advantages/Limitations
of Providing TAP (1)

+ quiet acquisition, threat not alarmed by radar illumination

+ increased coverage (range, azimuth, elevation)

+ warning of possible attack on AD fighter from outside own sensor coverage

+ protected (ECM resistance and MSG security)

+ improved IFF(friendlies equipped with MIDS/JTIDS and any information that may be supplied by these friendlies)

Figure 4a

Potential Advantages/Limitations
of Providing TAP(2)

+ reduced workload through visual presentation of flight info (also more rapid perception)
+ weapon firing coordination possibilities

- pilot must decide when and how to rely on externally provided data, quality of this data is very significant

    ü update rates of the external sensors may be inadequate

    ü tracking errors and intermittent data may produce correlation difficulties, missing and/or false targets

    ü RWR/RHAW indications may conflict with M/J information

Figure 4b

The augmented SILKA simulation software, at this time, could not cover all the above aspects due to time, manpower, and funding restrictions. Of primery interest was the effect of "quiet acquisition," i.e., approching the intruding aircraft in radio silence using the TAP information provided. A representative Central Region sub-scenario with 2 + 2 blue interceptor/fighters against a formation of 12 + 4 (12 fighter bombers + 4 escorts), with appropriate randomisation of starting positions was utilised. The scenario is derived from AC/243 (Panel X) RSG - 11 (Ref. 7) and related studies and is shown in Figure 5 giving a more detailed view of the combat area. Sufficiently large number of replications (more than fourty in each case) was carried out to permit the derivation of stable statistical results and many such runs were performed. More details are provided in Ref. 1.

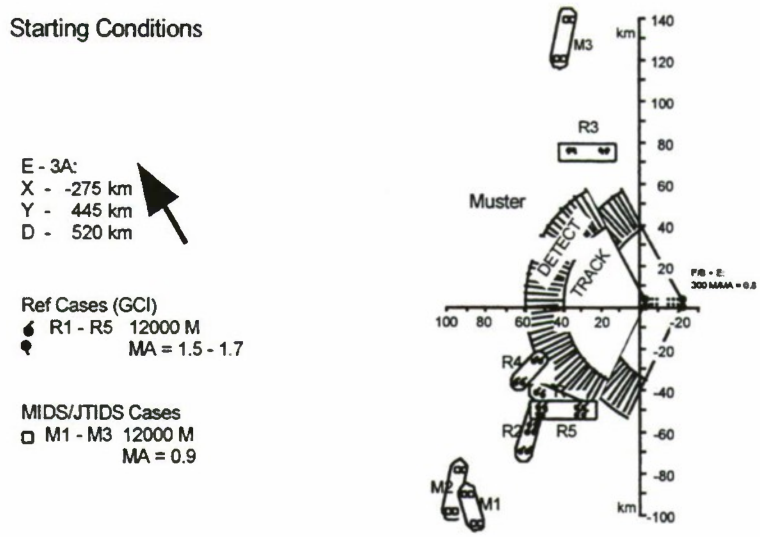


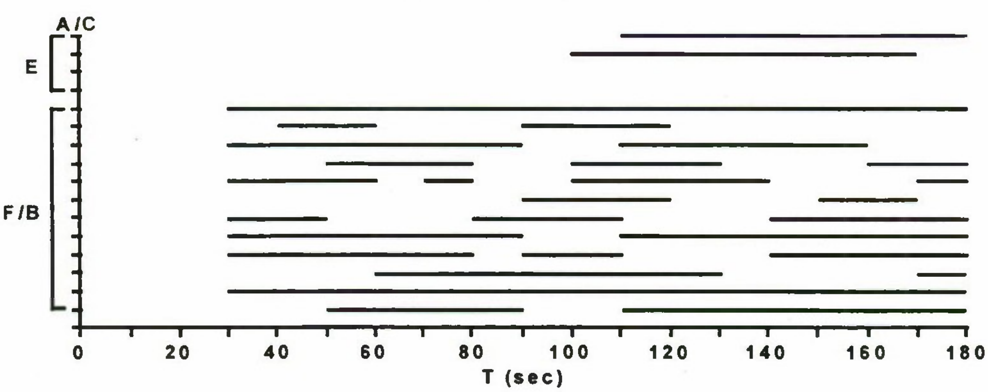
Figure 5

1.3 Simulation Results

Overall, within the framework of the scenario utilised and the assumptions made, the results indicated that significant improvements in effectiveness were possible when a TAP (including realistic errors, missing targets, (Figure 6)) is provided to the pilot of a modern fighter interceptor, with respect to the reference case A/GGI.

Track life, example

"BENIGN CASE"



Distance E-3A-Combat area - 500 km
RCS = 2.5 m ^{2} (Escort)
RCS = 5.0 m ^{2} (F/B)

Figure 6

A set of results comparing CAP (Case A) with A/GGI (Case B) as achieved during AC/243 (Panel X) RSG-11 study and comparing A/GGI (Case B) with MIDS (Case D) is presented in Figures 7a and 7b. The RSG-11 results were derived by using SILKA.

The MIDS related comparison indicates different time scales which consider the different approach strategies of the interceptor, e.g., the time difference of 130 sec represents the phase where the interceptor is approaching in the MIDS mode. Details can be found in Ref. 2.

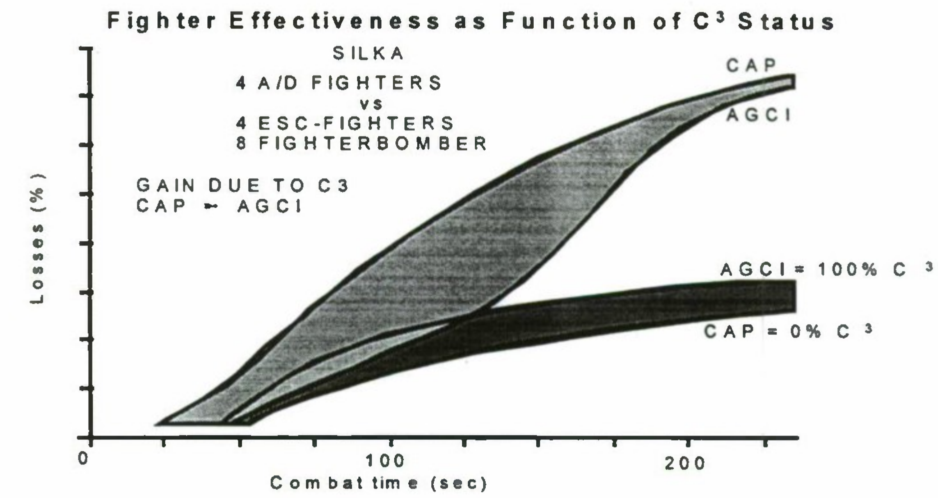


Figure 7a

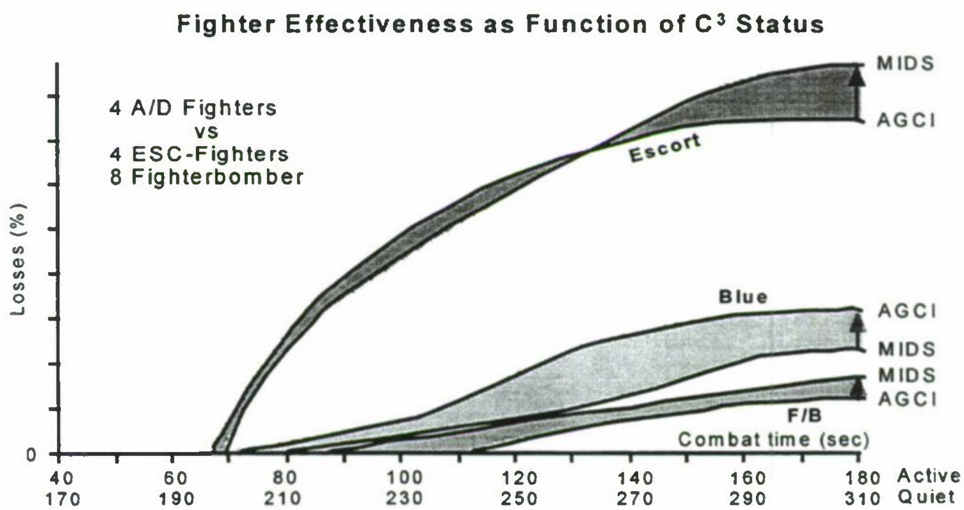


Figure 7b

Further results comparing the reference case with "MIDS case" (essentially, the provision of the TAP to the blue aircraft) imply an exchange ratio improvement of about 50% with respect to the reference case of NADGE A/GGI under perfect conditions.

The study was performed, as indicated earlier, under certain limitations. The modifications and improvements envisaged for the next study phase, as shown in Figure 8, will include large scale scenarios, multi-sensor tracking, identification, communication jamming, as well as, tactical implications on the fighters behaviour and the use of MIDS during combat.

Further Steps

- Extend air picture by implementing additional remote sensors
- Implement exchange of on-board information during combat
- Implement ESM
- Improve fighter tactics depending on improved information quality (outmaneouvre adversary's radar, coordinate firing, minimize radar search, etc.)
- Extend starting conditions

  ➽ Remark: Evolving in the present MIL-SILKA Activities

Figure 8

1.4 Conclusions

This paper has presented the results of a joint STG/IABG software simulation study on the contribution of a MIDS/JTIDS system to air combat. As far as is known, this is the first non-US activitiy of its kind. Although only a small subset of the multifaceted implications of MIDS/JTIDS could be studied, simulated, and many operational/tactics related questions were generated, the results indicated significant potential enhancements to air combat. The results are also significant from the viewpoint of focussing attention on new tactics and procedures which must inevitably derive from the availability of such a system. It is hoped to improve the knowledge of the MIDS related implications after the next study phase as discussed above.

References

(1) STC Technical Note TN-019, "MIDS in a Fighter/Interceptor," NATO Restricted, 1985.

(2) STC Technical Memorandum TM-811, "MIDS in Air Combat Simulation," NATO Confidential, 1986

(3) McDonnel Douglas McD-AS914, "F-15 JTIDS Definition Phase," 1979 (first three sequences of simulations).

(4) McDonnel Douglas McD-A6821, "F-15 JTIDS Definition Phase," 1981 (fourth sequence of simulations).

(5) Geiselhart, Ivey, Gavern, "A Simulation Study of JTIDS in Tactical Missions," CSDF ASD/ENECH Wright Patterson ADF Report, 1981.

(6) Wunderlich G., Braun R., "Air to Air Engagement Simulation," AGARD Conference proceedings No. 268 Modelling and Simulation of Avionics and C3 Systems, 1979.

(7) AC-243 (Panel X) RSG-11 Final Report.

(.) STC Technical Memorandum TM-669, "MIDS and UHF EW/ECM," NATO Secret, 1982 (See also STC IN-528.).

PART 2: THIRD PARTY TARGETING

2.1 Introduction

This study was performed by the 4 Power Senior National Representative Technical Group on Future BVR Missiles. The study objectives were as indicated in Figure 9:

Study Objectives

- To determine future air to air weapon systems to meet the threat
- To make a technological and systems recommendation for future air to air missile systems beyond those envisaged for the next generation (e. g., AMRAAM, ASRAAM, MICA)
- To identify subsystems and systems technology for future collaborative efforts

Figure 9

2.2 MISSILE CHARACTERISTICS

The Group started by defining the areas which appeared to be promising to be improved for future missiles (resp. are weak areas in present-day missiles (Figure 10)):

Future Air to Air Missile

- Longer range of engagement
  - ★ Propulsion
  - ★ Mid-course guidance
- Better terminal guidance
  - ★ Multiple targets
  - ★ Dense jamming
  - ★ Stealthy aircraft
  - ★ Target co-operation
- Advanced Seekers
  - ★ Improved IR
  - ★ Improved RF
  - ★ Multi-spectral
  - ★ Multi-mode

Figure 10

The major improvements appeared to be gained by improving range, seeker, and terminal guidance. Among others, the Group defined a missile concept, called SO2, wich had approximately the same mass as the reference missile SO1, a present-day missile, but had a ramjet motor and an Ku-band seeker. It is roughly described in Figure 11:

baseline performance
approx. 150 kg launch mass
active radar seeker (X band)
wings and control fins
rocket motor

2x baseline kin. range
approx. 150 kg launch mass
active radar seeker (Ku band)
wings and control fins
ramjet motor
external carriage combined with SRAAM

Figure 11

The performance is indicated in Figures 12 and 13. Figure 12 gives the increased fiving envelopes: In case of a non-manoeuvering target the range increase is about 35% (head-on) to 80% (tail); even more significant improvements are achieved in case of a manoeuvering target (3g): While the baseline missile looses about 60% of its range (head-on), SO2 indicates only marginal decreasing range performance.

Launch Success Zones

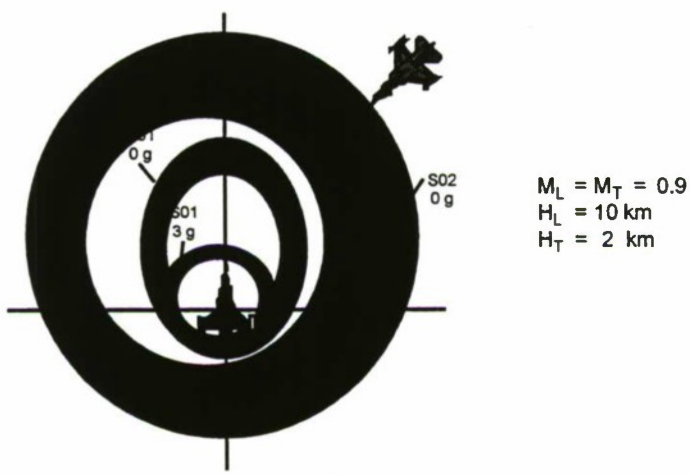


Figure 12

Concerning the seeker head, Figure 13 compares the various bands (x, Ku, Ka); it is shown, that Ku-band appears to be a reasonable option concerning its robustness in case of ECM; Ka-band would be best candidate for "benign" and "ECM", but worst case in case of rain which is a driving factor in European environment.

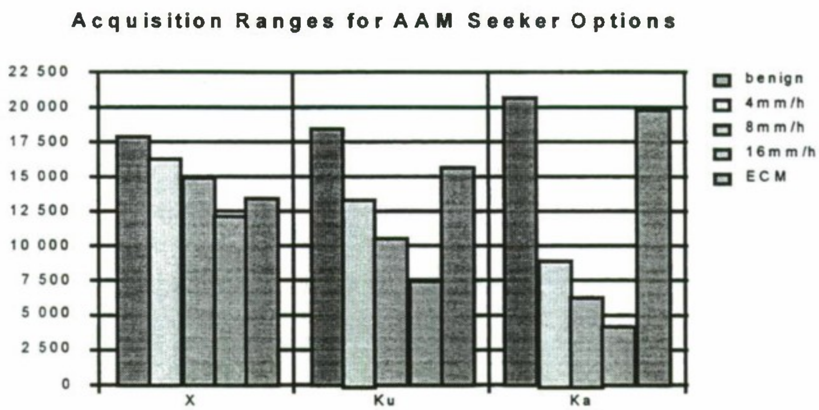


Figure 13

2.3 Third Party Targetting (TPT)

To make full use of the improved range of the SO2 concept, the possibility, and effectiveness of using Third Party Sensors was evaluated because the missile range by far exceeded the max. range of a fighter's onboard radar. It was decided to apply the NATO ACCS (Advanced Command and Control System) resp. their specification as indicated on Figure 14:

Third Party Targeting

- Offers large Pay-off for BVR Combat
- Required accuracy is difficult to achieve
- Single NAEW is not sufficient
- The best Track Accuracy Class (TAC) for non-manoeuvring targets assumed by NATO-ACCS:

Δx, Δy	:	km
Δz	:	km
Δvx, Δvy	:	km/s
track completeness	:	between % and %
range	:	up to km

Figure 14

This means that relatively crude information had to be used for missile launch and midcourse update because ACCS was not designed for this purpose. Nevertheless, this brought the study on the "safe side," because if TPT works in this (crude information) case it appears to be feasible.

A test was run to evaluate the JTIDS performance if being used as data link to update the missile. This test is presented on Figure 15.

Navigation Error TPT

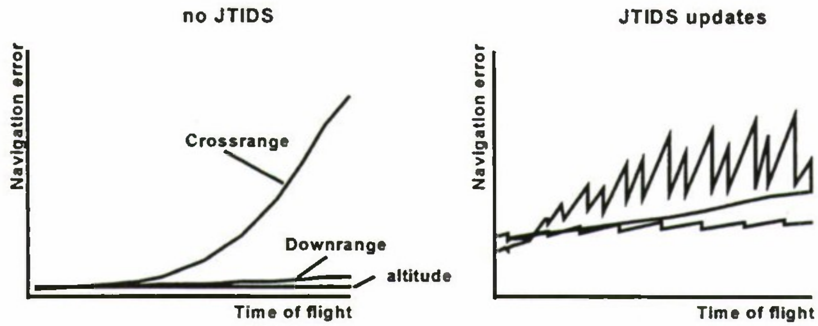


Figure 15

2.4 Simulation Environment

It was decided to apply a high density scenario which consisted of two air defence fighters against a force of eight fighter bombers covered by four escort fighters and two escort jammers (Figure 16).

Typical Scenario for Operational Analysis

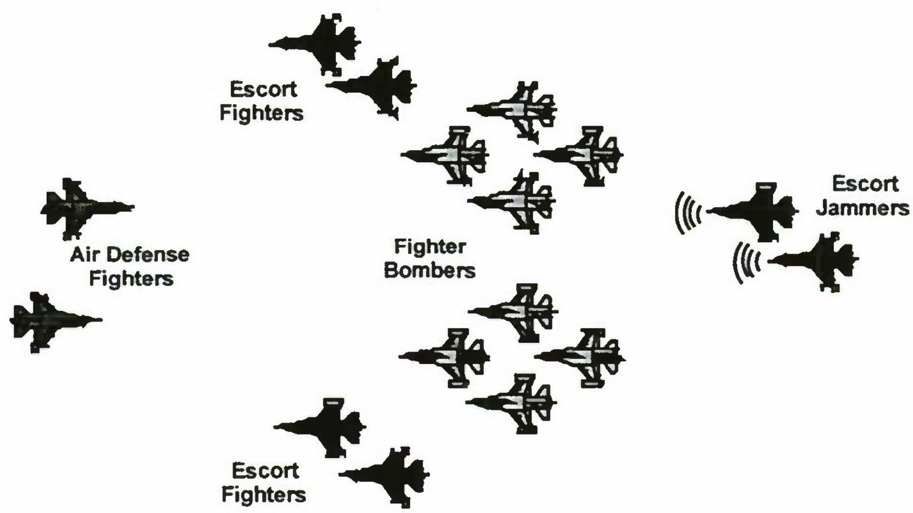


Figure 16

Furthermore it was decided to apply three different tools which were MIL-AASPEM (US), CEV (FR), and ARENA (UK). The entities of these tools are described in Figure 17 and indicate that they have quite differing feature, for example Workstations, Workstations + Cockpits, full digital.

Combat Simulations

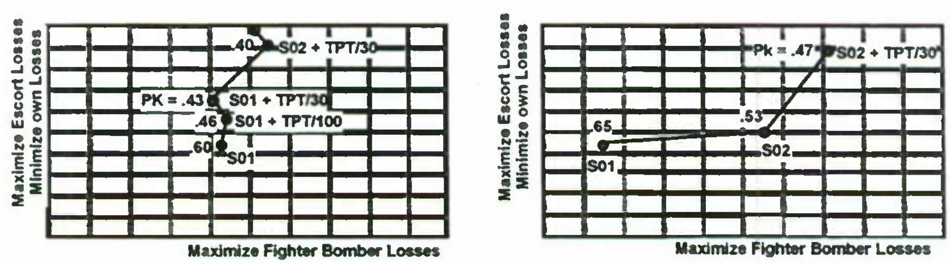
MIL-AASPEM	CEV	ARENA
2v2+4	2v4+8	2v2+4
Manned Workstations	2 Cockpits and Manned Workstations	Full Computer simulation
Medium altitude fighter versus escort	Medium altitude fighter versus escort	Medium altitude fighter low level escort
Bombers at low level	Bombers at low level	Bombers at low level
Free tactics for pilots	Free tactics for pilots	Pre-set tactics
No track degradation	Perfect Air Picture for Red	Realistically degraded Air Picture for Blue and Red

Figure 17

2.5 Simulation Results

Figure 18 presents the study results. Evaluation parameters are to maximize bomberlosses (x-axis) and to maximize the loss ratio escort-fighters / ad-fighters (y-axis), i.e., aim is to drive the result into the upper right corner of the figure. Figure 18 presents the impact of a "Wide Angle Search" (WAS) of the seekerhead; the seeker search volume had to be adapted to the ACCS information to prevent too much missile losses due to crude information. The left part of Figure 18 starts with the reference missile, SO1, roughly in the center of the figure, with a missile P_k of 0.60; applying TPT to this missile (for various degrees of air picture completeness) slightly increases the fighter exchange ratio in favour of the ad-fighters (but no improvement concerning the FB-losses); the P_k drops to around 0.45 (= increasing missile wastage). Applying the same TPT to SO2, a significant improvement in both fighter exchange ratio and bomber losses was achieved.

Effects of Missile Seeker, TPT



- No WAS
- TPT with independent RAP
- Fighter Co-operation

- Seeker with WAS
- TPT with independent RAP
- No Fighter Co-operation

Figure 18

The right part of Figure 18 implements both range improvement and WAS: SO1 (left from the center) to SO2 (= in the middle) presents the effect of range -improvement, and furthermore improving the result by adding WAS (= upper right); in any case of improvement, a slight to moderate decrease in P_k is to be observed, based on increasing fly-out time (= offering the target increased defence options) and increasing information failures from ACCS.

2.6 Conclusions

Figures 19 and 20 present the synthesis of the results from all three simulations.

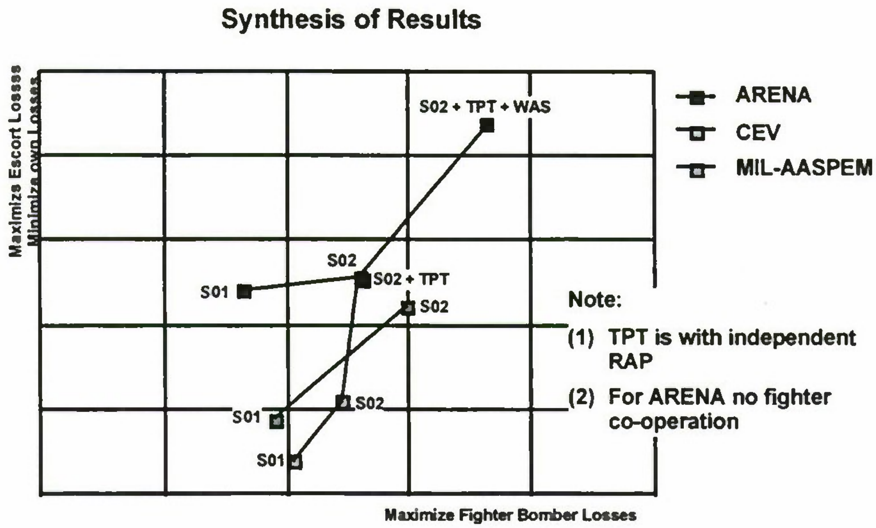


Figure 19

Synthesis of Results

- All 3 simulations show similar trends
- Bomber kills increase with use of S02
- Fighter attrition and fighter/escort exchange ratio high, with some exceptions
- Differences in results explainable by differences in scenario assumptions, tactics and subsystem modelling
- AI radar was a limiting factor in all 3 simulations
- Both CEV sim. and ARENA showed large gains in use of TPT with S02
- In some circumstances, Wide Angle Search in the missile seeker gave further improvements (ARENA)
- Further investigations of the balance between TPT and seeker performance are required

Figure 20

It should be mentioned that the results of this study had significant impact on the present FMRAAM activities in Europe, mainly the UK Staff Target for a future BVR missile.

References

4 PWR SNR (US, UK, FR, GE) Study (Post 1995 BVR AAM's)
Manned Combat Simulation Study/European Customer

ADAPTIVE AUTOMATION

** - Executive Summary only*

This page intentionally blank.

SMART COCKPIT CONTROLLER: THE USE OF ADAPTIVE AUTOMATION TO IMPROVE AIRCREW SURVIVABILITY AND MISSION EFFECTIVENESS

Norman W. Warner
Naval Air Warfare Center
Aircraft Division
Patuxent River, MD

Current crew stations include life support, escape and control/display sub-systems as separate entities within the aircraft system. Each sub-system operates along a rigid set of rules and without regard for the other sub-system's actions. In addition, each sub-system currently does not consider its effect on the aircrew's tolerance to mission stressors and the ability of the aircrew to process data. This lack of integration, coordination, and communication leads to a total crew station that provides either too little support, causing mission degradation, too much support which also leads to mission degradation due to discomfort and distraction or one sub-system action degrading the performance of another sub-system. This lack of integration and effective information management results in a loss of aircrew situational awareness and reduced mission effectiveness. The "missing link" is a Smart Cockpit Controller which effectively manages the functions of the life support, escape, and control/display sub-systems to achieve desired mission performance.

The objective of the Navy's Smart Cockpit Controller(SCC) project is to provide a system controller that effectively manages life support, escape, physiological support and cognitive/performance augmentation. The system will take into account the individual aircrew needs, past and present environmental stressors encountered, and escape scenarios. The mental workload demands on the aircrew will also be addressed in order to adapt the total crew station to the aircrew, thus enhancing mission performance. To successfully achieve the SCC objective, not only must the life support, escape and control/display sub-systems be effectively integrated but specific air vehicle sub-systems must also be considered and integrated. The following key crew station and air vehicle technologies will be integrated through the Smart Cockpit Controller: (1) Smart Aircrew Mission Support System(SAMSS) whose goal is to maintain homeostasis for maximum mission performance. Homeostasis is the maintenance of static conditions in the internal physiological environment regardless of the stressor presented. (2) Smart Escape System recognizes emergency, initiates ejection and optimizes seat function. (3) Smart Aircrew Interface between the SAMSS and Smart Escape systems which will incorporate adaptive automation and Helmet-Mounted Display technologies to provide relevant and timely information to the aircrew. (4) Advanced Flight Safety Systems including Terrain Avoidance and Auto Spin Recovery. (5) Vehicle Management Technologies which continuously assess the status of the aircraft sub-systems and reconfigure the aircraft to optimize performance. To integrate and effectively manage the information being transmitted between the above technologies, the core of the Smart Cockpit Controller will be adaptive automation. Adaptive automation is where the task allocation/information between machine and human is mutually shared in an adaptive manner. The invocation of the adaptive automation algorithms will be based on changes in the physical environment, crew station/air vehicle sub-systems, aircrew performance, aircrew physiology, or a combination. For empirical evaluation of the SCC, the Smart Cockpit Controller and the integrated technologies will be integrated into the F/A-18 Crew Station simulation within the Crew Station Technology Laboratory, Naval Air Warfare Center, Patuxent River, MD and the Dynamic Flight Simulator at the Veda's Warminster Facility. These evaluations will demonstrate the benefits of SCC and the integrated technologies on aircrew survivability and mission effectiveness.

The expected benefits of the SCC are (1) enhanced flight safety and survivability resulting in fewer aircraft losses, (2) reduced pilot workload and fatigue resulting in a reduction of aircrew deaths and injuries, (3) improved mission effectiveness due to effective management of crew station sub-systems, and (4) improved affordability due to increased aircrew and aircraft survivability. (Reprint of executive summary; formal paper not available)

This page intentionally blank.

COMBAT FLIGHT MANAGEMENT

Andrew Probert
Wright Laboratory
Vehicle-Pilot Integration Branch
Wright-Patterson AFB, OH

Summary: The future combat cockpit environment will differ from the past in two significant ways: 1) There will be much more information available to the pilot from off-board sources, and 2) Eye protection from frequency-agile lasers will block most or all light from entering the cockpit at times. Where operator task loading is already a problem, these changes will further increase pilot workload. Workload reduction for pilots, or remote operators of uninhabited combat aerial vehicles (UCAVs), is necessary if the benefits of information dominance are to be realized. Prior studies have shown a benefit to having: effective information management; automated in-flight mission planning, automatic trajectory generation, and auto-pilot coupling, as well as advanced pilot-vehicle interface technologies.[i,ii] A Combat Flight Management (CFM) system is being developed to reduce pilot or operator workload by integrating data fusion, correlation, and filtering with in-flight planning and flight control coupling; all imbedded in a unifying pilot-vehicle interface.

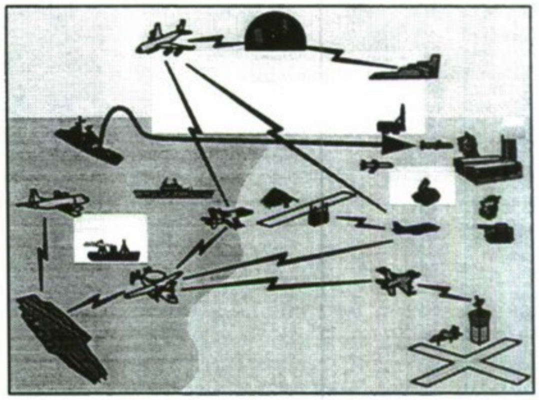


Figure 1. The System of Systems will bring information into the cockpit.

The Problem: Off-board information in the cockpit, provided by the networking of combat assets into a "System of Systems" (Fig. 1) will potentially improve situation awareness (SA) and help pilots locate targets, avoid threats, avoid adverse weather, avoid fratricide, and be re-directed by command and control elements to attack important time-critical targets. Unfortunately, the attention required to assimilate and assess the additional data, decide what to do, and to re-plan the mission accordingly, will add workload to already task-saturated cockpits (Fig. 2). Even if off-board data were effectively fused, correlated, filtered, and graphically displayed, giving the pilot perfect SA, there remains the problems of adjusting the mission plan to arrive at the target on time with enough fuel to return to base while minimizing threat exposure, and coordinating the changes with friendly forces.

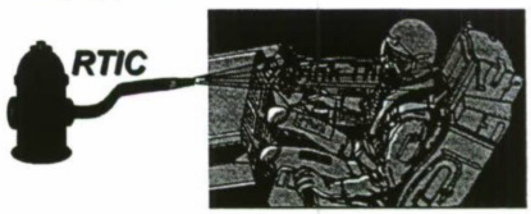


Figure 2. Information in the cockpit doesn't necessarily produce situation awareness.

Pilot workload will be increased because of the threat to pilot vision from frequency-agile lasers. Eye protection schemes will, at times, block most or all of the visible light spectrum, essentially turning day into night (Fig. 3). Since we already fly and fight at night, this is not a show-stopper; but, the lack of natural visual cues reduces SA and increases workload.

Figure 3. Eye protection from frequency-agile laser will turn day into night for the pilot, reducing SA and adding workload.

The result is that if nothing is done to reduce workload, pilots will be too task saturated to take advantage of the information dominance the US will enjoy. In a prior contracted effort with the Wright Laboratory's Flight Dynamics Directorate (WL/FI), called Mission Reconfigurable Cockpit (MRC), Lockheed Martin Tactical Aircraft Systems developed and evaluated future cockpit concepts for a generic single-seat multi-role fighter. In the first phase, Lockheed analyzed the challenges facing the designer of the future cockpit. In an analysis of nine generic functional tasks, workload estimations for projected future missions showed excessive levels for a single pilot during critical mission segments, using current technology, even without the presence of off-board data and the need to re-plan (Fig. 4). The top challenges (opportunities!) for pilot workload reduction were determined to be in: Information Management, Integrated Flight Control, Indirect Vision, Off-Board Data Integration, In-Flight Planning, and Target Cueing/Recognition. [iii]

Figure 4. Situation awareness is necessary but insufficient. Other tasks demand attention, precluding the preparation of alternative plans based on off-board data.

<u>The Solution</u>: After assessing projected technology capabilities, Lockheed designed and implemented a cockpit design concept in a development station which integrated several technologies to meet the identified challenges. Piloted simulation evaluations verified the reduced workload benefits of the concept.[iv] The Combat Flight Management (CFM) program will mature and integrate selected MRC technologies, thereby enabling full exploitation of off-board data (Fig. 5). Lockheed Martin will develop common core software modules applicable to existing and future platforms. The core software will be integrated into F-16 hardware-in-the-loop, piloted simulations for risk reduction, and for demonstration and evaluation of mission effectiveness and acceptable pilot workload. CFM is integrating technology in four areas:

1) On- and Off-Board Information Management: For the CFM program, Lockheed will take advantage of software developed in a contracted effort with the Wright Lab's Avionics Directorate (WL/AA), called Expanded Situation Awareness Insertion (ESAI). Given that ESAI will provide properly fused and correlated data, the display of information in the cockpit will be addressed within CFM by development of information management "policies." In this concept, even with perfect on-board/off-board data fusion, the amount of information available for display in a pilot's area of interest must be reduced to a manageable level by filtering track files based on various attributes of significance to the pilot. Several "policies" would be defined by the pilot pre-flight for selective task-oriented filtering of the pilot's tactical displays. The pilot would be able to manually control the display "policies" in-flight.

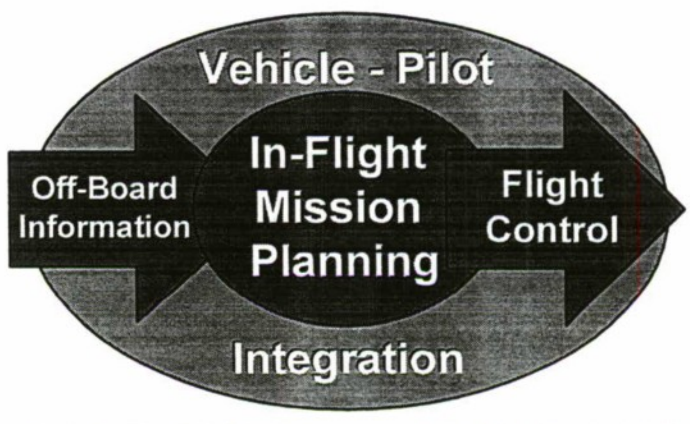


Figure 5. Combat Flight Management will enable tactical air assets to exploit information dominance by reducing workload for pilots or remote operators.

Concepts for display formats and control mechanizations developed and evaluated in the MRC contract will be adapted for the F-16 simulation demonstration.[v] These concepts include plan-view map formats, symbology, and switchology.

2) In-Flight Mission Planning: Lockheed Martin Aeronautical Systems, Marietta, GA, has been working on in-flight planning for some time. DARPA's Pilot's Associate program developed a planner for a generic high altitude, low observable, air superiority mission. For MRC, Lockheed adapted it to low-level operations by including terrain masking and intervisibility.[vi] An in-flight mission planner would begin a mission with the planning data and routes carried forward from a ground based planning station like the Air Force Mission Support System (AFMSS). As the mission progresses, the planner assesses the situation based on updates from off-board sources, system status, and pilot input. The planner would automatically propose a new route for pilot acceptance if: 1) A new target location is designated, 2) New threats impact the current plan, 3) The pilot deviates from the current plan, or 4) The pilot directs a change or update to the current plan. The planner would consider the same signature and threat data, and guidance from the air tasking order (ATO) (special instructions, rules of engagement, etc.), flight manuals, etc., that the pilot and AFMSS does in pre-flight planning. The planner's priorities would be: 1) Execute the ATO (e.g., attack the target on time); 2) Have enough fuel to return; 3) Minimize threat exposure (or, more restrictive for special operations, minimize detection); and 4) Minimize fuel use. The planner would not estimate survivability, judge the prudence of continuance, nor recommend aborts. An abort plan will be continuously maintained, however, in the event the pilot wants it. The pilot would be free to modify the plan in any way by: inserting or moving "must fly" steer points, targets, air refuelings, or delays; directing sensor use plans; changing the recovery base; or setting speeds, altitudes, or courses.

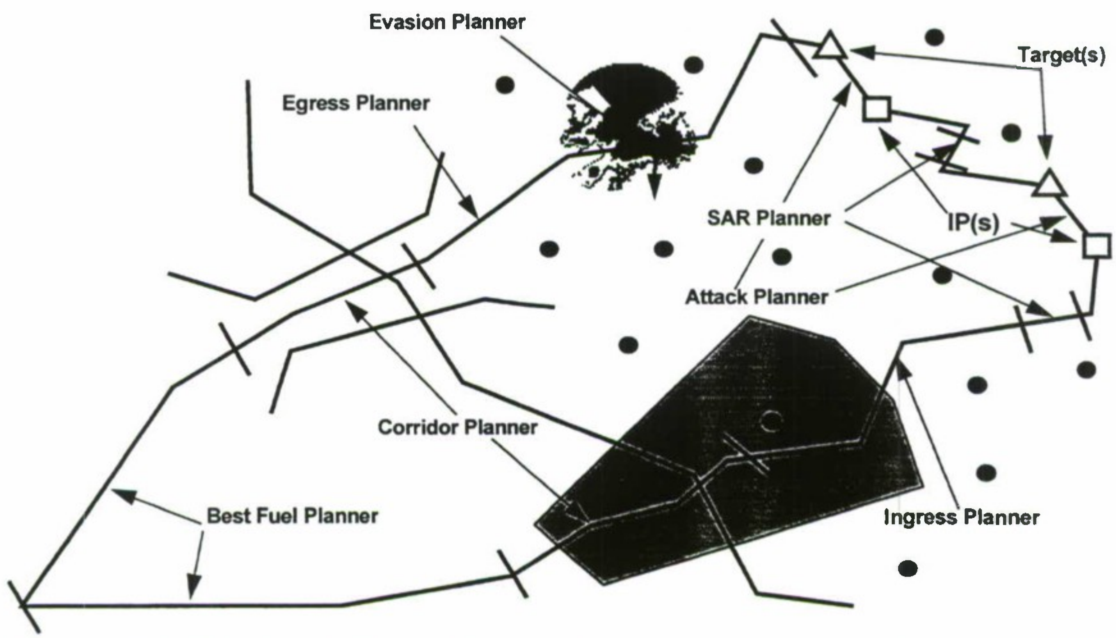


Figure 6. The CFM In-Flight Mission Planner will construct entire mission plans using specialized planners for separate mission segments.

Figure 6 shows how the planner would construct an entire mission plan using different planners for separate mission segments. Specialized planner functions for particular applications include sensor use trajectories, attack plans, threat avoidance navigation, detection avoidance navigation, and minimum fuel use. The diagram also shows the planner complying with safe passage procedures with the corridor planner. The evasion planner is a quick reaction mode to a threat that popped up within lethal range without prior warning. The planner would find the safest escape path, then would re-calculate a plan to continue the remainder of the mission.

Without an in-flight planning capability, the pilot's only secure option when faced with off-board information on a new threat may be to abort the mission. If he attempted to circumnavigate the threat area, he could risk running out of fuel. Attempts to penetrate a threat area without an optimized trajectory could increase risk of threat engagement. MRC simulations evaluated survivability benefits of automatic in-flight terrain-masking threat avoidance route planning.

In 1995, 72 sorties were flown in piloted simulations in the MRC development station at Lockheed Martin in Fort Worth.[vii] While attempting to penetrate Surface to Air Missile (SAM) defenses at low altitude, 36 sorties were flown with automatic in-flight planning, and 36 without. With a moderate (day 5) threat density ingress to a target area, pilots were alerted to pop-up SAM threats by simulated off-board assets. With automatic planning available, the system generated proposed trajectories for the pilot's acceptance that minimized exposure to the threats. Without automatic planning, pilots manually inserted steer-points into their route with a cursor and graphic interface on a map display. The trajectories flown in the virtual simulation were recorded and fed into a constructive threat simulation (SUPRESSOR) for survivability analysis. Use of the automatic planner reduced the average threat radar track duration to less than one fifth of that for the manual re-planning cases (Fig. 7). Interestingly, the auto planner had more total threat tracks. However, because of the short duration, there were no shot opportunities (Fig. 8). The use of manual planning resulted in several shots and two kills in the 36 sorties.

	Manual Planning	**Auto Planning**
Avg. Track Duration (seconds)	**25.6**	**4.8**
Avg. Mission Duration (seconds)	**643**	**614**
Percent of Mission	**4.0**	**0.8**

Figure 7. Survivability Impact - Average threat radar track time reduced 81% with automatic vs manual planning. (From SUPRESSOR.)

	Manual Planning	**Auto Planning**
TRACKS	**9**	**15**
SHOTS	**12**	**0**
KILLS	**2**	**0**

Figure 8. As a result of lower average track time, there were no shot opportunities with auto planning.

The subjective results of the simulations were also positive. Pilots were asked to rate the workload involved when new threats were introduced. Using the Modified Cooper-Harper scale, where workload is rated from 1 to 10 (1-easy, 10-impossible, over 3 is unacceptable), pilots favored the automatic planner (Fig. 9). The pilots' opinions on the concept were also favorable (Fig. 10).

3) Integrated Flight Control: The most significant task in the pilot workload equation is that of flying the aircraft. As workload increases beyond a pilot's capabilities, a successful pilot will employ a strategy of task shedding that will maximize survival. *If manual flight control is required, it must be the last to be shed, after all others including target acquisition, weapon delivery, defensive reactions, or SA maintenance.*[3] Flight control automation may be a cheaper and more effective means to reduce workload than other efforts such as automatic target recognition. If the pilot can rely on the flight control system to safely follow the trajectories generated by the planners, more attention can be paid to the study of sensor

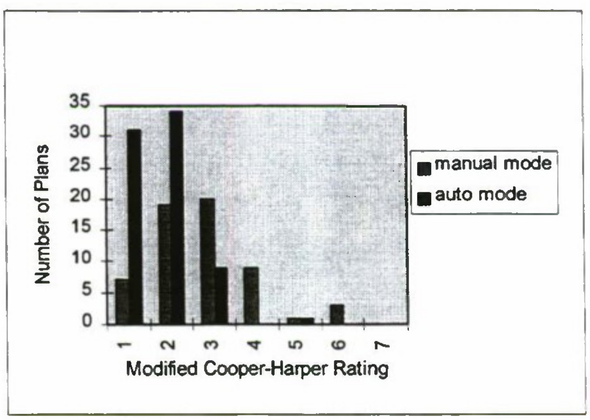


Figure 9. Workload for route replan with pop-up threat.

imagery, and aimpoint selection. More of the pilot's attention will be available for defensive SA and force package coordination when things don't go exactly according to plan. The integrated flight control concept was part of the MRC simulation evaluations. Pilot opinions support the concept (Fig. 11).

For integrated flight control automation to be reliable enough for the pilot to trust and use it, the planned routes and trajectories must be compatible with the capabilities of the flight control system. For example, it would make no sense if the planner calculated a terrain masking threat avoidance route below altitudes the aircraft's terrain following system was capable of flying. In addition to automatic coupling to navigation routes, the flight controls must integrate with weapon delivery (fire control), sensor use, and defensive end-game trajectories, if available.

- "It takes the mental gymnastics out of the equation."
- "Outstanding aid to overall mission success."
- "...Essential in making future aircraft more effective and decreasing vulnerability. The ability for the MP to incorporate TOT, fuel management, and threat assessment is outstanding. A definite plus to our future combat pilot."
- "A significant workload reducer."
- "It doesn't help the war to bailout due to fuel starvation rather than getting shot down."

Figure 10. Pilot comments on automatic in-flight planning capability from MRC simulations.

- "A lot of head-down time [is required] to maximize weapon/avionics effectiveness - Mission Integrated Flight Control [is] critical to shift pilot workload from aviator to manager."
- "Significantly reduces pilot workload and is essential to prevent task saturation."
- "Excellent"
- "As our requirements for survivability, flexibility and responsiveness increase with the continuing explosion of the sensors (on/off board) information, pilot workload must be significantly reduced. If I can fly the whole mission in the simulator with my hands in my lap, then in combat, the workload will be manageable, that is, the fog and the friction of war will become [a] stressor I can devote full attention to."

Figure 11. Pilot comments on integrated flight control from MRC simulations.

There will be times when the pilot has no planned trajectory to follow and needs to maneuver, for whatever reason (formation positioning, target search, hazard avoidance, etc.), while engaged in other attention-demanding tasks. A low workload means of directing the aircraft must then be provided. A Pilot-Directed Guidance (PDG) mode automatically maintains speed and altitude (above ground level or mean sea level), yet responds to steering commands.[viii] PDG inputs via the rudder pedals would allow the pilot to steer while handling various mission equipment such as maps, charts, flashlights, cameras, binoculars, visors, HMDs, personal life-support gear, food, piddle packs, etc. A PDG mode for remote operation of a UCAV would aid system flexibility and responsiveness.

4) Pilot-Vehicle Interface (PVI): Because CFM is not meant for any one particular system, there cannot be dependence on any particular cockpit display layout or device. Each System Program Office or prime contractor will have to develop the exact display formats and control mechanizations. This task should not be slighted, because the success of the integration will depend on it. Commercial airlines' experience with flight management has shown that poor integration can result in pilot confusion and accidents.[ix]

In CFM, Lockheed will demonstrate the concept in an F-16 cockpit. Lockheed will incorporate PVI technologies that will make it possible for new channels of communication to be used between the pilot and the aircraft. Display media include helmet mounted displays (HMDs) and 3-D audio. Speech recognition technology will be applied to reduce the number of switch actions required and allow pilots to make inputs without diverting attention from visual tasks (Fig. 12).

- Voice Recognition
  - "Concept is great!"
  - "Cuts down on switch hits when reliable."
  - "Enhances the PVI tremendously. An innovation that is long overdue."
  - "Impressed."
  - "...it would seriously reduce mundane house keeping chores, reduce # of switches, complexity, and increase RM&S."
- Helmet Mounted Display
  - "HMD is too heavy, cumbersome, and would provide a significant source of revenue for optometrists. It is, however, one of the single most critical technologies in this program and must be developed further."

Figure 12. Pilot comments on PVI.

These PVI technologies are less dependent on specific cockpit geometry and could be retro-fit in other systems without undue expense, provided the processing power is available. Their incorporation may, in fact, result in cost savings. For example, transfer of other functions beyond flight management to the voice recognition system could allow the removal of hardware such as the up-front control keypad and instrument panel buttons and switches. This, in turn, could remove ergonomic constraints and make the accommodation of smaller pilots easier and cheaper.

Since the planner would require the presence of a terrain database for intervisibility calculations and route generation, the HMD could easily be used to display a computer rendered image of that database with unlimited field of regard for off-axis imaging. Such a system was implemented and evaluated in an F-16 simulation in 1994.[x] Pilots experienced with LANTIRN were asked to evaluate the effectiveness of the Synthetic Terrain Image (STI) against LANTIRN. STI was rated as being more effective at providing terrain awareness for maneuvering at low altitude in rugged terrain, especially for high G defensive maneuvers. Objectively, pilots were able to fly lower with STI. In addition to off-axis viewing, STI would provide an image when an infrared (IR) system would be blanked by poor thermal contrast or visibility. STI in a HMD would provide a remote UCAV operator situation awareness that may not be practical with IR sensors.

Conclusion: Prior work indicates there is much to be gained if flight management concepts become reality, for both piloted and "uninhabited" combat aircraft. Flight management, however, is not so much technology development as it is technology integration of the traditional technology core areas of avionics, flight control, artificial intelligence, and human-factors. The Combat Flight Management program intends to provide the necessary risk reduction and evaluation effort that will allow users to fully realize the potential flexibility, lethality, and survivability that the Information Warfare revolution promises.

Acknowledgments: Ray Hanson, the Lockheed Martin Tactical Aircraft Systems program manager for MRC and CFM; the MRC lead engineers: Gene Dawson, Dick Spengler, and Dr. Randy Aust; Tony Leavitt of Lockheed Martin Aeronautical Systems, the developer of the in-flight planning system; and Dr. Aust, again, for editorial assistance on this paper.

REFERENCES:

[i] G.R. Rundle, et al, *Flight Management for Air-to-Surface Weapon Delivery,* General Dynamics, Fort Worth, TX, AFWAL-TR-84-3118, Feb 85.

[ii] G.N. Maroon, et al, *Tactical Flight Management Exploratory Development Program,* McDonnell Aircraft Company, St. Louis, MO, AFWAL-TR-85-3027, Mar 85.

[iii] G.C. Dawson, et al, *Mission Reconfigurable Cockpit (MRC) Human Engineering Progress Report, Vols. 1 and 2,* Lockheed Fort Worth Co., Fort Worth, TX, WL-TR-94-3010 and 3011, Nov 93.

[iv] G.C. Dawson, et al, *Mission Reconfigurable Cockpit (MRC) Phase 2 Report, Vol. 1: Human Engineering Progress Report (HEPR),* Lockheed Fort Worth Co., Fort Worth TX, WL-TR-96-3062, Nov 95.

[v] R.P. Spengler, *Mission Reconfigurable Cockpit (MRC), Phase 2 Report, Vol. 2, Human Engineering Design Approach Document-Operator,* Lockheed Fort Worth Co., Fort Worth TX, WL-TR-96-3063, Oct 95.

[vi] C.A. Leavitt, "Real-Time In-Flight Planning," *Proceedings of the IEEE National Aerospace and Electronics Conference (NAECON '96),* p 83-89.

[vii] R.B. Aust, "Mission Planner Evaluation in the Mission Reconfigurable Cockpit," *Proceedings of the IEEE National Aerospace and Electronics Conference (NAECON '96),* p 102-108.

[viii] R.A. Coppenbarger, "Helmet Mounted Display Symbology for Automated Nap-of-the-Earth Rotorcraft Flight," in *Helmet- and Head-Mounted Displays and Symbology Design Requirements,* Ronald J. Lewandowski, Wendell Stephens, Loran A. Haworth, Editors, Proc. SPIE 2218, 351-360 (1994).

[ix] D. Hughes, et al, "Accidents Direct Focus on Cockpit Automation," *Aviation Week and Space Technology, 30 Jan 95, 52-60.*

[x] W.H. Corwin, et al, *Synthetic Terrain Imagery for Helmet Mounted Display, Vols. 1-3,* Honeywell Technology Center, WL-TR-95-3025-7, Nov 94.

A UNIFIED WEAPON SYSTEM: THE ROLE OF ADAPTIVE AUTOMATION

Estrella M. Forster
Naval Air Warfare Center Aircraft Division
Patuxent River, Maryland

Introduction. Intelligent Systems are currently being introduced into the Tactical Air Environment. Significant advances have been made on escape, life support, and aircraft recovery systems. As advances in technology take place, scientists and engineers have identified the need to address the impact of such advances on the components of the weapon system: the operator, the cockpit, the aircraft, and, ultimately, the mission performance of the overall weapon system.

That is, as the machine is becoming independently "smart," the research community recognizes, and the operator demands that the progress in technology be coordinated with the human component of the same. However, there is a lack of data addressing this relationship. Meanwhile, sponsors of scientific research are necessarily interested in ensuring their particular requirements directly and positively benefit the fleet, the customer. To this end, three issues have come to the forefront when addressing the support of technology advances which will ultimately provide the pilot with the premium equipment and materials to complete his or her mission: affordability, integration, and transition.

A key aspect of this support is integration. To achieve the integration of the aircraft, the cockpit, and the operator, a "dialog" between the various programs and research disciplines is indispensable. Such dialog would lead to an investment group which has a common goal: an advanced and integrated weapon system which maintains the operator in the loop. To this end, efforts have been initiated to converge the technology advances of several ongoing programs into a <u>unified weapon system</u>. The Adaptive Automation concept may be considered the conduit to such a system.

Background. Currently, the weapon system is partitioned into the human operator and the machine it operates. The various technologies and information emerging to improve this system are either: a) integrated by the operator or b) integrated into the system independent of the operator. Neither solution is appropriate. Having the pilot integrate the information increases workload and, automating the systems for the pilot, transforms the pilot into a "monitor;" an out-of-the-loop entity which probably will not accept new technology over which she or he has no significant input. Instead, what is required, is to bring an integrated system to the operator. Adaptive Automation serves to accomplish this goal. According to the American Heritage Dictionary, *adaptive* refers to "...something, such as a device or mechanism, that is changed or changes so as to become suitable to a new or special application or situation. According to the same text, *automation* refers to "acting or operating in a manner essentially independent of external influence or control." Hence, Adaptive Automation, also called Adaptive Function Allocation, implies a "division of labor" between machine intelligence and the operator. Adaptive refers to the change in responsibility for performing a function(s). The process of altering this responsibility is Adaptive Automation.

There is a lack of operational data addressing: a) the impact of Adaptive Automation on the mission environment, b) the integration of the person and the vehicle this person controls (and wants to remain controlling), c) the metrics assessing the person-machine relationship, and d) the implementation of Adaptive Automation concepts. Hence, there are several approaches to address the development of adaptive automation algorithms. In general, this control logic attempts to address: 1) the environment, 2) human performance, and 3) human physiology. Ideally, the final control algorithm would include all three elements.

Developing Adaptive Automation. Developing adaptive automation requires addressing several issues including assessing a) what information is significant, b) how to monitor that information, and c) how to make a decision based on this information to invoke the allocation of functions both to the machine or the operator.

<u>Environment</u>. Issues concerning the environment include: a) the aircraft status (e.g., altitude, attitude, G history); b) the cockpit status (e.g., stick input, information display); and c) the mission environment whether it is combat, emergency, or routine. All three issues require the pilot in the loop.

Human Performance. Issues concerning performance take into account both motor and cognitive performance. Examples of this element are: a) assessing operator performance under stress (e.g., altitude, acceleration, thermal); b) assessing performance under various mission scenarios in both static and dynamic simulators; and c) assessing the "teaming" arrangement between pilot and machine.

Human Physiology. Issues concerning human physiology include monitoring the electroencephalogram (EEG), the electrocardiogram (ECG), the electrooculogram (EOG), etc., in an effort to assess pilot status.

Towards Adaptive Automation. One program addressing all three Adaptive Automation concerns (environment, performance, and physiology) is the Smart Aircrew Integrated Life Support System (SAILSS). The development of SAILSS goes hand in hand with the development of Adaptive Automation and its implementation in future aircraft. The objective of this effort is to develop navy requirements and demonstrate an integrated systems approach to a man-mounted aircrew personal protective ensemble which provides biofeedback control to regulate system parameters. SAILSS will:

- Provide G protection.
- Protect against the "Push/Pull" effect (negative to positive Gz).
- Provide altitude protection.
- Maintain body core temperature.
- Maintain skin temperature.
- Eliminate dehydration due to suit effects.
- Provide early warning of threats: chemical, biological, G-induced loss of consciousness (G-LOC), and other Altered States of Awareness (ASA).

SAILSS may be thought of as a an "undergarment" with physiologic sensors (e.g., EEG, ECG, etc.). "Smart" indicates the use of biofeedback with which the system will reduce the direct effects of environmental stresses on the aircrew's physiologic and cognitive state. Also, SAILSS takes into account aircraft status (e.g., altitude, attitude, G history). After signal conditioning, the various signals will be analyzed by a computer. This computer will control anti-G suit pressure, breathing gas pressure, breathing gas mixture, body thermal control air flow, and other support and protection functions. Ultimately, SAILSS will maintain physiologic homeostasis by providing real-time adjustments of life support equipment, optimizing cardio-respiratory and brain function, and improving mission effectiveness.

SAILSS serves as an example of how Adaptive Automation can serve as a conduit to a unified weapon system where SAILSS provides "...automation that changes as a function of changes that occur during the operation of the weapon system."

Investment Group. Current major research and development programs addressing crew and aircraft systems are generating intelligent systems that will significantly impact the mission of the weapon system and the safety of both the aircraft and the operator. These technologies have similar concerns:

1. HOW should the technology be implemented(hardware, software, interfaces)?
2. WHO is going to invoke the technology (pilot, technology itself)?
3. WHAT is going to invoke the technology (critical events, mechanisms)?
4. WHEN is the technology going to be invoked (timing, frequency)?
5. WHY is the technology going to be invoked (thresholds, critical events)?
6. WHERE should the information regarding that technology be displayed (location, format)?

Similarly, Adaptive Automation concepts encompass concerns in that it involves the "identification of the functions to be allocated and the identification of the parameters upon which a function is to be allocated." Ultimately, these concerns, and the solutions to the same, directly impact the development of other technologies and their integration to form a unified weapon system. Therefore, to achieve this unified system, we need to initiate or strengthen current dialogs among the various programs. Such dialog would lead to an investment group which has a common goal: an integrated weapon system which maintains the operator in the loop. The common factor among the technologies is the pilot. This pilot needs and demands to be in the loop. As we integrate the various systems, this loop needs to be defined. Adaptive Automation plays a key role in defining this loop in that it addresses the environment of the overall weapon system and the performance and physiology of the operator.

HELMET-MOUNTED DISPLAY CONCEPTS, SITUATIONAL AWARENESS, AND ADAPTIVE AUTOMATION

Scott D. Fechtig, Eric E. Geiselman, Norman W. Warner, and Harold D. Warner
Naval Air Warfare Center Aircraft Division
Patuxent River, MD

Introduction: The potential for a head-coupled helmet-mounted display (HMD) system to aid pilot situational awareness (SA) in the tactical air environment is significant. This is due primarily to two major display features: First, critical information is available to the pilot via a see-through combiner regardless line-of-sight or head movement. Secondly, as a function of the head tracking transducer, information can be presented overlaid on the outside world thereby making it spatially compatible with the human's natural processing of the surrounding environment. These features act together to reduce the visual area the pilot must scan to gather information from otherwise dispersed display sources. Empirical evaluations suggest that HMD presented information affords the pilot more time with a line-of-sight closer to the primary area-of-interest in the absence of any associated performance decrement. A possible interpretation of this finding is that the HMD produces increased SA by enabling the pilot to prioritize visual attention without the burden of routine information gathering activities.

Current Program: Ultimately, the value of the HMD as a SA aid depends on how effectively the information presented by the system is formatted, mechanized, and fused. In this context, the various forms of information include text, symbology, sensor imagery, and augmented reality applications (e.g. synthetic terrain). As an effort to optimize the information presentation of the HMD, the NAWCAD Crew Systems Division Human Performance Technology Branch has initiated a program to investigate mission-based functionality of the HMD across near, mid, and far term technology maturation time frames. The program approach includes mission decomposition and task analysis, concept development, dynamic simulation, man-in-the-loop evaluation, and production of a unique multimedia desktop-based dynamic concept visualization tool. Program objectives include evaluating advanced information presentation concepts on the basis of objective performance, workload, and situational awareness.

Adaptive Automation: Integration of the HMD with adaptive automation technologies and concepts will be considered for the mid and far term maturation time frames. As a display device, the HMD is uniquely capable of multiple functionality by acting as both a system input component as well as a means of completing the feedback loop by providing system and/or situation status to the pilot. For example, the HMD head tracker provides a continuous estimate of pilot line-of-sight and head movement to the ownship system. Simultaneously, the system has access to the information being processed and displayed to the pilot and knowledge of the nature of the area-of-interest (why the pilot is looking in that direction). These variables can be fed logically into automation routines to help prioritize handling of other aircraft functions such as automated defensive counter-measures, target sorting, etc. Also, as the pilot concentrates visual resources on the area-of-interest, information displayed by the system will display status and keep the pilot aware of automation state changes.

Future Research: We anticipate the HMD functionality concepts developed during this program will reveal several areas where further research is indicated. One likely area is the functionality of the coordinate reference frame produced by the use of the head tracker technology. How best to use (and avoid the use of) this new information representation capability is an important area of empirical investigation.

(Reprint of executive summary; formal paper not available)

This page intentionally blank.

USE OF SIMULATION AND MODELING
FOR ADAPTIVE AUTOMATION CONCEPTS

Richard E. Edwards, Doug Swanson
Boeing Defense & Space Group
Seattle, Washington

J. Yarmi
Microvision, Inc.
Seattle, Washington

Introduction: Situation awareness within the integrated battlespace depends upon a rapid and accurate assessment of the spatial relationships and operational states among many dynamic players. Current tactical aircraft use two dimensional displays to present this information, forcing operators to create a three-dinensional mental model of the mission environment by integrating selected information from several two-dimensional displays. The mental effort required to perform this task during time critical portions of a mission increases cognitive task loading and degrades mission performance. Application of emerging three-dimensional display technologies to tactical situation displays could reduce, or eliminate, these performance decrements by presenting complex spatial relationships among geographically dispersed warfighters in a natural format that is immediately usable.

Virtual Retinal Display: One of the promising display technologies, which appears to satisfy many of the mission and crew requirements, is the Virtual Retinal Display (VRD). This device can be used to present full-color, high-resolution (SVGA), binocular images containing text, graphics, or video data. The display may be "see-through" or immersive, and display transparency can be manipulated on a pixel by pixel basis. Display brightness can be adjusted to compensate for the high ambient light levels of fighter cockpits and for dimly lit conditions such as night flying. These features make the VRD a good candidate for application of real-time adaptive automation concepts to address information presentation and control requirements within rapidly changing mission environments. In addition, the VRD can be integrated with a helmet for fighter applications, or with light-weight goggles or spectacles for helicopters and Command, Control, Communications and Intelligence (C^3I) platforms. Aircrew safety is improved through the use of light-weight VRD components that reduce pilot fatigue and neck loads during ejection and through the use of low voltage DC power rather than the high voltages needed to drive cathode ray tube displays.

(Reprint of executive summary; formal paper not available)

This page intentionally blank.

TASK ALLOCATION BETWEEN REAL AND VIRTUAL PILOTS PUTTING REAL INTELLIGENCE INTO A VIRTUAL COCKPIT

Maxwell J. Wells and Hunter G. Hoffman
Human Interface Technology Laboratory (HITL)
University of Washington
Seattle, Washington

Recently Wells, Baldis & Hoffman (1996) demonstrated the concept that crewmembers in different geographical locations can interact with one another in a virtual cockpit. Two-player teams assumed the role of pilot and copilot in a shared virtual cockpit. Virtual representations of each person (avatars) were used to communicate with voice and gestures as the crewmembers collaborated on a task requiring teamwork.

'Telesavance' refers to the transmission of the state of a person's 'wisdom' ('savance' comes from the noun 'savant' meaning a wise or knowledgeable person) or situation awareness using telecommunication techniques. We hypothesize that an advisor (e.g., an AWACS operator) could be more efficient at communicating Situational Awareness (SA) if they could simultaneously assess the state of the pilot's SA (their savance). With face-to-face communication this SA transmission is done by sending information, and by sending probes ('OK?', 'Know what I mean?',) and by looking for verbal and gestural confirmation that the information is understood, or not. Segal (1994) has shown that much communication in the cockpit is nonverbal. According to Segal, being able to observe the activities of your crewmate (e.g., where they are looking, pointing, switching, if they are nodding, etc.) provides valuable information about how well they understand the situation, their intentions, and the status or mode of the equipment they are manipulating. This paper explores the concept of using avatars and shared virtual environments to replace face-to-face communication.

Telesavance represents a unifying construct, some potential advantages of which are: 1) As a design heuristic it will push toward solutions that emphasize the communication of situation awareness, rather than the communication of data; 2) Because it uses natural intuitive communication techniques, it may require less cognitive resources; 3) It can serve as in interim step on the road to an intelligent assistant in the cockpit; 4) It requires less bandwidth than video (73,728,000 bits/sec for video vs. 11,520 bits/sec per sensor); 5) Gestures can be filtered and used in machine recognition; and 6) Flexibility of information presentation (e.g., see-through avatars).

Domains in which telesavance might play a role include 1) Networked UAVs: Spatially separated operators of UAVs could be networked so as to enhance the performance of complex missions requiring close coordination of assets; 2) Team training; 3) Task allocation between onboard and remotely located pilots; and 4) The integration of automation into the cockpit, using a virtual crewmember to represent the machine. We see this latter domain as an interim step on the road to implementing adaptive automation.

Our initial experiment, reported in the presentation, was a test of the utility of savance. A pilot simultaneously controlled two Uninhabited-Combat-Air-Vehicles (UCAVs) on two separate missions with the help of an advisor. The experiment explored whether allowing the advisor to see and hear the pilot enhanced the communication of SA, compared to verbal-only communication. Performance assessment included mission success, team interaction, communication patterns, and evaluation of the players' SA.

(Reprint of executive summary; formal paper not available)

This page intentionally blank.

UNINHABITED COMBAT AIR VEHICLES: HOW MUCH AUTOMATION IS TOO MUCH?

JOSEPH C. VON HOLLE, Lt Col, USAF
Joint Cockpit Office
Wright Patterson AFB, OH

ABSTRACT

The United States (US) Department of Defense (DoD) is committed to increasing the roles of uninhabited air vehicles (UAVs) beyond use as reconnaissance platforms. In the future, UAVs may be used as a force multiplier in many scenarios -- even as lethal weapon systems. There are many areas to be examined for accurate situation awareness for the warfighters to take advantage of this new weapons platform opportunity. The range of choices from fully automated (e.g., cruise missiles) to fully human controlled (e.g., remotely piloted vehicles) comes with its own range of benefits, limitations, and costs. The trade off between human control and automation takes center stage in this debate.

It is time to begin thinking ahead and asking the hard questions about the application of Uninhabited Combat Air Vehicles (UCAVs) in future combat environments. More questions than answers will be presented in this paper. This paper will examine two areas critical to the ability of the combat crews to maintain mission situation awareness for UCAVs.

First, it will examine the question of automation and autonomy. The benefits, limitations, and costs must be examined on the continuum from fully piloted to fully automated. The amount of autonomy may vary at different mission phases. Greater autonomy for the vehicle reduces the need for direct human interaction but increased the need for high speed processors and memory on the vehicle. More direct human intervention drives communication requirements such as band width, speed and security. Through it all, the human operators, whether acting as "pilots" or mission managers, must have thorough situation awareness provided by the sensors and displays. The criteria developed in such an examination will enable developers to find an appropriate balance for the environment in which the vehicle will operate.

Second, the physical distance of several hundreds of miles between the human operator and the vehicle creates a set of user interface problems which require in-depth analyses. The physical distance will impact mission decision making and decision timeliness as well as crisis management when the vehicle malfunctions or encounters hazards. Likewise, the physical and psychological effects of the time lag, maneuvers without motion cues, long-term use of synthetic vision/virtual environments, etc. must be examined. Some studies have been performed but the overall knowledge set is still inadequate.

The US Air Force and US Navy have begun research in areas that should provide solutions to many of the questions. These studies include Expanded Situation Insertion, UAV Human Interface, Real Time Information Fusion, Flight Control Automation, and multi-ship control.

INTRODUCTION

The concept of UAVs is not new. However, the increased emphasis gives the research and development community of the US DoD an opportunity to develop, test, and acquire these vehicles through a process which results in optimum capabilities at an affordable price. The analytical assignment of functions and tasks to humans and machines through the process is critical to achieving system optimization.

This paper presents a short background of UAVs, a brief survey of what is achievable in flight control automation, and a discussion of the issues of providing optimal situation awareness for uninhabited <u>combat</u> air vehicles, capable of delivering lethal munitions.

This paper, while well researched, contains personal opinions and projections. The opinions are those of the author and do not represent the official position of the US Air Force.

A BRIEF HISTORY OF THE UAV

In many ways, UAVs are now where manned aircraft were in about 1915... used primarily for forward observation and artillery spotting for ground forces. The path forward also looks like the history of

manned flight as we move toward a concept of lethal weapon delivery platforms. However, the time to grow to lethal capability will be compressed from 80 years to about 10-12 years.

The use of remotely piloted vehicles (RPVs), now often used as a synonym for UAVs, is by no means "new" in the 1990s. Records of aviation history indicate remotely piloted vehicles were used for testing of aerodynamics and controls as early as the 1920s. In 1937, the Navy completed and flew a conversion of an N2C2 trainer to a radio controlled drone (Figure 1) for what is often cited as the first US RPV. The concept was eventually adapted for use as an assault drone during World War II; the first use of a "cruise missile" by the United States.

Figure 1

By the 1960s, numerous contractors were building the vehicles which were the forerunners of today's drones, decoys, and cruise missiles. The vehicles were most often used in photo reconnaissance roles including flights over China and Vietnam following the Soviet shoot-down of a USAF U-2 in 1960. Vehicle developers tried to apply the idea to autonomous vehicles for "seek and attack" missions against enemy radar sites but automation of that time was not up to the task. However, the concept of attack vehicles was still tested during the 1970s when the USAF tried delivering Maverick missiles from a modified reconnaissance drone during the Vietnam war (Figure 2). That concept proved imperfect but the ground work of the future was in place.

Figure 2

Research continued in the 1970s as the concepts were refined. The USAF tested the Compass COPE B, a high altitude, high endurance vehicle, in 1973. The Compass COPE gave birth to the idea of vehicles like today's Global Hawk. Unlike most previous RPVs, which are ground- or air-launched and retrieved in mid-air, Compass COPE was designed to take off and land from conventional runways. Among the tests was a successful endurance flight of 17 hours 24 minutes at altitudes of more than 55,000 feet. The Compass COPE program moved through a series of successes and failures before the Air Force decided against a production version.

Appreciation of the value of UAVs for U.S. military, however, did not emerge until their use during Operations Desert Shield and Desert Storm. Prior to that, operations in Grenada and Libya identified the need for an inexpensive, unmanned, over-the-horizon targeting, and reconnaissance for force commanders. The Navy acquired the first operational system since Vietnam when they bought the Pioneer RPV in the late 1980s to support Marine Corps operations. They later modified it for shipboard deployment. The Army received its first Pioneer system in 1990. During Desert Storm, with 85% of the US manned tactical reconnaissance assets committed, UAVs emerged as a "must have" capability. Six Pioneer systems in all (three with Marines, two on Navy ships, and one with the Army) participated. They provided highly valued near-real-time reconnaissance, surveillance and target acquisition (RSTA) and BDA, day and night.

Today's UAVs are categorized as either tactical (relatively short range and endurance), such as the Pioneer (Figure 3) or endurance (high altitude, long range, and high endurance) like the Global Hawk

(Figure 4). The tactical UAVs satisfy the requirement for surveillance and reconnaissance capabilities to ensure force protection, enhance situation awareness, and increase the effectiveness of organic fire support systems. The endurance UAVs will satisfy the joint force and theater commanders' needs for broad-area coverage with options for continuous dwell. A mix of tactical and endurance UAVs, coupled with manned and space-based platforms, will provide capabilities that expand the commanders' potential to dominate the battlespace.

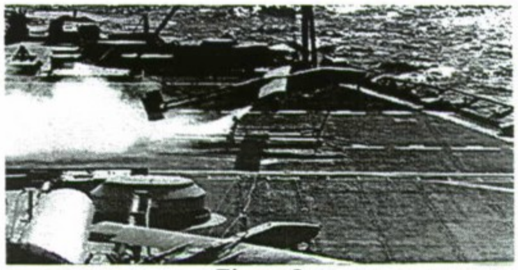


Figure 3

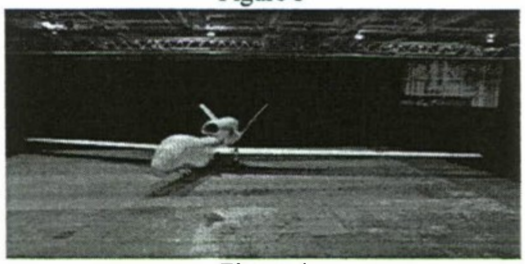


Figure 4

The US military is already exploring concepts for lethal UAVs. In 1995, the Air Force commissioned both its Scientific Advisory Board (SAB) and the Air University (AU) staff to project how the US air forces might operate 25-30 years in the future. The results of both studies envisioned UAVs performing combat roles with or in place of manned platforms. (Figure 5 shows the Scientific Advisory Board's conception of an Uninhabited Combat Air Vehicle.)

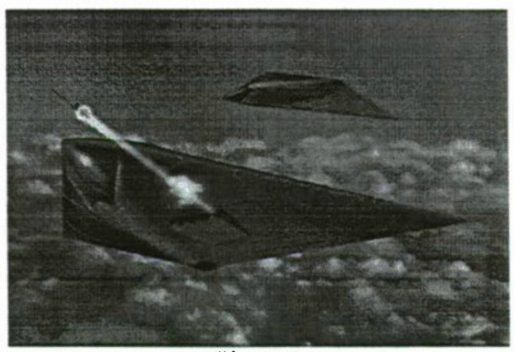


Figure 5

AUTOMATION - WHAT'S ACHIEVABLE

Like UAVs themselves, automation is not new. The concept of automation for flight systems began with the idea of imbedded computers in aircraft systems. Several advances were made in the 1960s, especially in automated flight control, which led to insertion of high-level capabilities during the 1970s and 1980s on many civil and military aircraft. Here are examples of advances in the past two decades:

- The Lockheed L-1011 airliner flew virtually hands-off across the ocean from throttle up to power off more than 25 years ago.

- The F-16 Falcon must be under constant automated flight control to remain in straight and level flight.
- The F-117 Nighthawk flies a complete mission from wheels up to wheels down without pilot input except for weapons release.
- The F-22 Raptor pilot will be an "Executive Systems Director," controlling high level mission functions while the aircraft flies.

Still, for the foreseeable future, there will be very few cases where a military aviation mission will be accomplished entirely by an autonomous automated system. Humans, with their ability to make judgments with incomplete information or in the face of unexpected circumstances, remain an essential element for safe, reliable, and effective mission performance. The future for UCAVs will likely be a mix of automated systems and human interaction; a shared responsibility for mission accomplishment.

There are many approaches to automation, typically dealing with a balance of what is technologically achievable and what is desirable for human knowledge and control of the mission. In the New World Vistas report, the Air Force SAB expressed it this way,

> *"Perhaps the most critical issue pacing the evolution of UAVs is that of manual versus automatic control of the wide range of functions to be executed during a mission. Human controllers have limitations (such as the number of parameters that can be controlled simultaneously and the speed at which humans can respond to sensed changes), but they also have unique abilities not yet replicated in automatic controllers. The human can learn to perform control functions and can thus adapt to unexpected inputs and demands. Humans can also reason effectively under conditions of uncertainty and perform higher order integration tasks."*

The SAB's New World Vistas study cited multiple levels of automation for UCAVs. These levels can be summarized into three groupings: Highly Autonomous; Human Supervised; and Human Controlled (Figure 6).

The first describes the highly autonomous vehicle that will be fully automated for flight control and programmed to make "human-like" decisions based on established flight rules, planned mission requirements, and circumstances. For the highly autonomous vehicle, the human acts as the "battle commander" providing high level inputs to perform activities such as: search sectors, identify possible targets, initiate targeting sequence, and report for authorization.

Second, is the human supervised vehicle. This vehicle would still contain highly automated systems for flight control, target search and identification, and image relay. However, the human acts as the mission commander directing the vehicle and sensors to certain search areas while setting parameters such as waypoints, speed, and altitude.

The final level of automation requires a good deal of human intervention. The vehicle automation is restricted to flight control limits and some navigation. These vehicles, much like today's tactical UAVs, are under constant control of the humans including some "hands-on" flying.

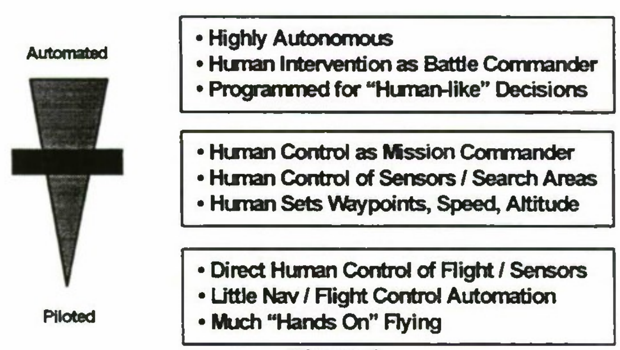


Figure 6

Decisions must be made about the appropriate level of automation for a UCAV based on many factors including vehicle parameters, mission requirements, and available technologies. The allocation of tasks to automation will impact the entire system. The decision to commit more of the functions to automation may reduce the crew size required to operate the UCAV and reduce the display fidelity and temporal

fidelity (closeness to "real time") requirements but will certainly increase the required on-board processing capability and drive up the software complexity. Conversely, allocating more of the functions to the human would allow a reduction in on-board processing and automation software complexity but may increase the complexity of hardware and software at the control station.

These decisions also impact the required bandwidth for the system. When the UCAV is highly automated, little demand is placed on the system to transmit the control information. This can free bandwidth for information from the vehicle, such as sensor imagery, or enable the vehicle to remain more autonomous and less "observable" through radio transmissions. The reverse applies when the human is more involved in vehicle control. More bandwidth is required for control which either reduces information from the vehicle or drives an overall increase in system bandwidth requirements.

The past tendency has been to exploit available technology for automation and leave the human the remaining tasks (plus any new tasks that may be generated). The failure of this approach is to consider the automation opportunities and the human as parts of a larger system and to allocate functions in a way to optimize the system.

SITUATION AWARENESS IN A UCAV

The necessity of good situation awareness cannot be overstated... and what applies to manned strike aircraft applies to UCAVs as well. The underlying moral code of the United States and the public's reluctance to accept US military and non-combatant casualties creates a need for human involvement in activities which result in the delivery of lethal force. So, the principle for developing a UCAV, as eloquently stated by Colonel James Francis of the Defense Advanced Research Program Agency, is to keep the pilot's head in the aircraft but leave the rest of him at home. To accomplish this goal, the human controllers and the automated systems must be able to react to the dynamic changes of the combat environment. And, their ability to react depends heavily on how well they can 'see,' 'hear,' and 'feel' what is happening around the UCAV.

A typical strike mission consists of a series of mission phases which can be roughly sorted as take off, ingress, attack sequence, egress, and landing (Figure 7). Automation for a UCAV performing a strike mission must be adaptable to the various phases. For example, the ingress phase may require the vehicle to be "stealthy" and limit radio transmissions. At this time, the vehicle may operate in a highly autonomous mode for flight control and defensive tactics. However, the attack sequence and egress may require more human intervention for target identification, weapon release authority, damage assessment, re-attack, or defensive evasion.

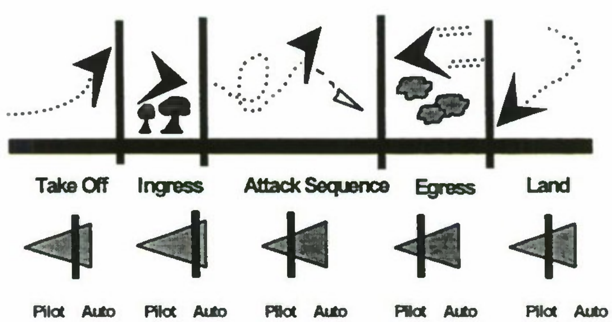


Figure 7

The decision about the most appropriate level of automation for each phase of the mission must be made based on many factors including mission requirements, information needs, and the capabilities of the humans and automation. Good situation awareness is based, to a large extent, on the appropriate information being available at the appropriate time. The strike mission must be carefully analyzed to ascertain the information required as well as the sequence and timing of the information delivery. Since the mission involves release of lethal munitions, a final, critical set of information for release decision must be determined.

The physical movement of the pilot from the air vehicle to an alternative control location introduces a number of issues which must be addressed to enable good situation awareness. Foremost among these issues is the need to replicate the tactical environment of the vehicle for the physically separated pilot. Research is required to determine the methods and develop the capabilities for the pilots to apply their flight knowledge, intuition, and "gut" reactions when they are up to 200 nautical miles away from their vehicle.

Central to this issue is the pilot's ability to "see," "hear," and "feel" the environment the vehicle is experiencing. The level of situation awareness at the pilot end will equate to positive situation control at the vehicle end. If the pilot will be provide a limited visual field of view, alternative display methods must be used to "fill in" the missing information. On the other hand, if the decision is to provide a synthetic environment, the developers must understand the impact on the technical requirements to achieve that environment. The level of automation applied to functions in each mission phase must be considered to assure the system supports the information needs of the human. This is largely technology dependent but operational decisions must be made as well.

Another key factor beyond just replicating the pilot's awareness is the information and automation capabilities required to recover from enemy jamming, system failures, or battle damage. Each of these situations presents a special problem for vehicle awareness and control from a remote location.

The tactical strike mission is an information intensive environment. But the information must be available in a timely fashion or it becomes useless. This highlights another issue for recreating the environment of the tactical strike pilot. The tactical information must be passed over great distances and through several processors. The distance and the processors can create bottlenecks that impede the timeliness of the information. Most often the manifestation of the bottlenecks will be a time delay in the signal between when the sensor "sees" something and when information is presented on the controller's displays. These delays determine how close to "real time" the human will operate.

There are numerous locations of information handling that may contribute to the bottleneck. The transmission of data is dependent on the bandwidth of the data link, transmit and receive rates of the radios, and whether data must be encoded or encrypted prior to transmission. Any information required from other sensors (not on the UCAV) may impact the data transmission delays. Also, the physical distance, especially if satellite communications are involved, and the signal attenuation can be factors in the information bottleneck.

The next source of bottlenecks or delays is the vehicle mounted and ground based processors. The processor speed, memory type and memory device speeds, bus architectures, and display capabilities are all factors in the information handling speeds. Likewise, the automated functions and the speed of the automation can impact (positively or negatively) the speed of information processing.

The information bottlenecks described can have a compounding effect but, for the most part, the smallest bandwidth at any one point becomes the maximum throughput for the system. The current information transmission network, conceptually represented in Figure 8, is reliable, robust, and capable. However, it has not been tested for its ability to support UCAV operations. Nor, perhaps, have the UCAV concept developers considered all the impacts of the technologies that make up the network.

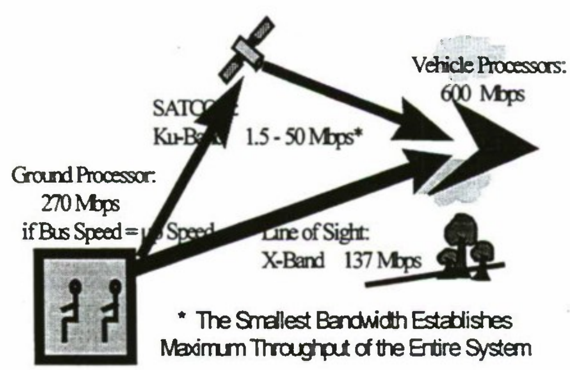


Figure 8

The final node in the information processing network is the human operator. The human, between the displays and controls, is typically expected to use their interpretation of displayed information to determine appropriate control actions. The combined effect of human cognitive processing time and technology/network delays could be a major player in the success or failure of UCAV operations.

The problems caused by the combined effects reveal themselves in a number of fashions. The first manifestation may be misjudgment of position, speed, distance, or attitude. This could cause increased temporal disorientation between the vehicle and the operator leading to gross misjudgment of position, visual acquisition errors, and pilot induced oscillations. The human's effort to compensate could significantly increase the cognitive task workload and eventually result in poor or incomplete situation or risk assessment and operational mistakes. The downward spiral in performance would continue as operator frustration and fatigue become a factor.

The bottom line is these combined effects can cause very poor situation awareness. An obvious solution is more automation. This can lessen the human operator's workload and move the processing closer to the "real time" sensor information. The problem in the past has been our approach to automation. As already stated, the past approach has been to take advantage of what is technologically feasible and work the human's role back into the equation. A close look at mishap reports shows the outcome of this approach. The cause is rarely due to automation failure but rather the failure of information management and communication between the machine and the operator.

It is time (again) for a hard look at how we develop systems. The failures of the past have been an inability to design automation that fits naturally into a human organization. Methods still are required to describe how the characteristics, capabilities, and limitations of the human partner fit into the system. Likewise, we need to be able to describe and analyze how the interaction of the human and machine impact the larger "system of systems" into which they must fit. This requires a deeper understanding of human crew functions, decision making, and teamwork under stressful conditions.

Just as important is to do a better job of designing the automation with the human partner in mind. This requires intelligent and adaptive automation that clearly communicates its status, decisions, and actions to the human partner. The automation, depending on the system design, should be subservient to a human commander at some level. Finally, the automation must be reliable and robust to assure appropriate decisions and actions when the communication link between the human and machine is broken.

This should be achievable. But, it requires much work from both the developers of automation and the human factors community. The solution should begin with realistic function allocation... taking advantage of the best capabilities of both the human and the automation. The human must always have meaningful involvement and not be maintained merely in the role of a system monitor. Likewise, the automated system must be designed to provide understandable and traceable decision paths so the human can react if the automation is unavailable. The human factors community and the automation designers must develop a better environment of communication and trust to finally develop a capable and coherent system.

In the end, we must remember the UCAV will deliver lethal weapons. The one place where automation always fails is in accountability. Humans must always be involved in activities or processes which result in the delivery of lethal force.

First, we, the developers of future UCAVs, must analyze the UAV philosophy from the ground up for UCAVs. Clearly, there are valuable lessons to be learned from the UAVs which are already flying. But the UCAVs and their application will be so different as to justify not accepting the current state of the art in UAVs, carte blanc, as the starting point. They must be treated in their own fashion and developed from a clean sheet.

Second, we need to develop effective methods and tools for human-system function allocation. The old tendencies described earlier must be set aside to assure an appropriate partnership and shared responsibility between the human and the machine. This must begin with a clear focus on the capabilities and limitations of both in determining their roles in the final system.

Third, remember that the UCAV will be integrated into an existing system of manned and unmanned assets. This requires a "system of systems" design approach to make a full determination of the UCAVs performance requirements. These requirements will guide the development of the human-machine interface. We must strive to develop an atmosphere of quality teamwork throughout the system. The new system must be designed to fit easily into the larger environment without requiring the larger environment to undergo significant changes to accept the UCAVs.

Finally, we must be sure to involve the operational warfighters now and throughout the process. The knowledge and insight of the operational forces should be serve as a guide and standard for the development of the UCAVs. We must develop the systems to meet the situation awareness demands of the warfighters. To the developers, situation awareness is something to be analyzed, measured, and tested... to a warfighter it often means whether or not they will see their family again.

References

- United States Air Force Scientific Advisory Board, *New World Vistas: Air and Space Power for the 21st Century - Summary Volume,* 15 Dec 95

- United States Air Force Scientific Advisory Board, *UAV Study Report*, SAF/PA 96-1204 UAV, 6 Dec 96

- Advisory Group for Aerospace Research and Development (AGARD) Conference Proceedings #520, *Combat Automation for Airborne Weapon Systems: Man/Machine Interface Trends and Technologies*, April 1993

- Defense Airborne Reconnaissance Office (DARO), *UAV Annual Report - FY 1996*, 6 Nov 96

- United States Air Force Museum Archives, via the US Air Force Museum homepage at: http://www.wpafb.af.mil/museum/

INVESTIGATION OF ECGs FOR USE IN ADAPTIVE AUTOMATION

Harold D. Warner
Naval Air Warfare Center
Aircraft Division
Patuxent River, MD

Norman Warner
Naval Air Warfare Center
Aircraft Division
Patuxent River, MD

The U.S. Navy is currently exploring the use of adaptive cockpit automation as a means of ameliorating the adverse effects of high workload exposure on pilot performance and situational awareness during military flight operations. As envisioned by the Navy, various cockpit functions will be transferred from manual pilot control to automatic control when pilot workload reaches a critical level. A fundamental issue in the automation of these functions concerns the extent to which pilot workload can be quantified <u>objectively</u>. Recently, a technique called the Measurement of Adaptation Processes (MAP) was developed for measuring changes in the physiological state of the pilot using electrocardiographic (ECG) data. This technique was subjected to a two-phase flight simulator assessment to determine its value as a measure of pilot workload. The present paper describes the MAP technique and the results of the simulation.

<u>MAP Technique</u>: The MAP technique was developed in the private sector by a team with expertise in science, engineering, and medicine. In the application of the MAP technique, ECG data collected during flight are subjected to post-flight power spectrum analyses. These analyses produce power spectra as a function of time whose amplitude, according to the system designers, represents the combination of heart rate and respiration rate. It was felt that the MAP technique might be useful in quantifying pilot workload for application in adaptive automation because of the changes in heart rate and respiration rate that occur concomitantly with changes in workload.

<u>Flight Simulator Assessment I</u>: The first test of the MAP technique for measuring pilot workload was conducted using a static flight simulator configured to represent the advanced F-22 fighter cockpit. Computer-generated imagery was used to provide a highly detailed out-the-cockpit visual environment. A tactical combat task was performed by the study participants comprised of terrain following as wingman in a two-ship formation, two types of bomb deliveries, and air-to-air combat against four adversary aircraft. Analyses of the ECG data collected during the flights revealed marked increases in the amplitudes of the power spectra at different points in the task. The power spectra, however, were not time-coded with sufficient precision to determine whether the amplitude changes actually occurred as a function of increases in pilot workload.

<u>Flight Simulator Assessment II</u>: Because the MAP technique showed promise as a candidate for implementation in adaptive automation, a second simulator test was performed using an improved version of the MAP system software that provided more precise time coding of the power spectra. This made it possible to statistically test the relationship between changes in amplitude and workload. An F/A-18C flight simulator with computer-generated external cockpit visual imagery served as the test platform in this investigation. The subjects' task was to fly as wingman in a two-ship formation flight. ECG data were recorded during the flights, and RMS slant range between the aircraft served as the performance measure. The difficulty of the formation task was varied to produce five levels of pilot workload, as follows:

- Workload Condition 1. Straight and level flight.
- Workload Condition 2. Gradual level turns.
- Workload Condition 3. Frequent altitude changes.
- Workload Condition 4. Frequent turns and altitude changes.
- Workload Condition 5. Maximum rate turns and altitude changes exceeding the pilots' ability to maintain formation with the lead aircraft.

Three experienced pilots served at participants. Each participant was tested under each of the five workload conditions, and there were three repetitions of each condition. The ECG data collected during the simulator flights were subjected to post-flight spectrum analyses using the improved MAP technique. Correlation coefficients between power spectrum amplitude and RMS slant range were then obtained for each of the workload conditions. The results of these correlational analyses are discussed.
(Reprint of executive summary; formal paper not available)

This page intentionally blank.

DECISION AIDING

* - Executive Summary only

This page intentionally blank.

DESIGNING TO AID DECISIONS THROUGH SITUATION AWARENESS ENHANCEMENT

Mica R. Endsley
Texas Tech University
Lubbock, Texas

Stephen J. Selcon
Defence Evaluation & Research Agency
Farnborough, United Kingdom

Introduction

Automation and various forms of artificial intelligence have been the focus of concerted efforts in cockpit development over the past two decades. Work in this area has largely been focused on the problems of high workload in the cockpit — either as an affect of reduced crew size or increased complexity of missions and avionics systems. In addition, these systems typically strive to increase the effectiveness and reliability of operator decision making.

Success in these endeavors has frequently been difficult to achieve, however. This is partially due to the underlying difficulty of creating such systems. Eliciting complex decision behavior from pilots (or other experts) and modeling that knowledge into a sufficiently robust system has proven extremely challenging for a variety of technical reasons and due to the complexity and manner of human intelligence. Even more significantly, however, increasing evidence suggests that attempts to improve human performance through the addition of traditional expert system-like decision aids may not lead to the desired benefits (Endsley & Kaber, 1996; Endsley & Kiris, 1995; Kibbe, 1988; Selcon, 1990) . In general, if system advice is ambiguous or incorrect, the system is far more likely to reduce human decision quality and speed. Evidence suggests that decision support system advice represents a new source of information that must be combined with other information, thus adding to the decision problem as much as assisting it. The introduction of automation has also been shown to lead to significant problems in trust, complacency, and loss of situation awareness (SA) that can lead to out-of-the-loop performance problems when the human is called upon to compensate for automation deficiencies or failures (Endsley & Kiris, 1995) . Furthermore, evidence suggests that these systems do not necessarily reduce workload, but rather introduce new forms of workload (Hart & Sheridan, 1984; Wiener, 1988; Wiener, 1993) .

This does not mean that efforts to improve performance through decision support are in vain, however, but rather that such efforts may have been directed at the wrong problem. Due to the difficulties in combining decisions from human operators and machines and due to reductions in SA that have been found with automated systems, our approach is to seek to improve decision making by focusing on improving SA. Most such systems have been focused on making decisions or performing tasks automatically. Trained operators are quite adept at making decisions, however, once they have the correct SA. The majority of their problem is in getting and maintaining high levels of SA. Over 88% of aviation accidents involving human error have been attributed to problems with SA (Endsley, 1995b) . Therefore, systems which help operators to achieve high levels of SA are most likely achieve the desired decision making and overall performance improvements.

Investigation

An investigation of a system for directly presenting needed SA to fighter pilots was conducted. The study examined the use of displayed threat envelopes (as shown in Figure 1) to provide pilots with needed information for replanning around pop-up threats (rather than a system which provides a recommended route, for instance). Such "explanatory displays" have been shown to be a highly effective means of providing decision support, improving both performance and trust in the system (Fletcher, Shanks, & Selcon, 1996; Selcon, Bunting, Coxell, Lal, & Dudfield, 1995a; Selcon, Smith, Bunting, Irving, & Coxell, 1995b) . In the study, pilots were provided a part-task mission simulation. The simulator consisted of a cockpit mock-up with a head-down display, an out-of-the-cockpit world display with a HUD overlay, and a stick and throttle. The pilots' task consisted of following a course of waypoints in a low level ingress flight to a target at 450 knots and 2000 feet altitude. Pop-up threats appeared during the flight which pilots were required to avoid, staying as close to their pre-determined flight path as possible. Three aircraft were presented as pop-up threats at the subject's altitude for approximately 30 seconds. Twelve experienced RAF pilots served as subjects during the experiment.

A within-subjects experimental design was used in which each subject was exposed to eight trials in each of two conditions: (1) no envelopes - in which only the threat aircraft symbol and supporting information (aircraft type, speed) were shown; (2) with envelopes - in which the same information was supplemented with a direct graphical display of the threat aircrafts' launch success zone (LSZ), as shown in Figure 1. These LSZ envelopes were dynamic based on algorithms that took into account aircraft type, speed, location, and relative aspect.

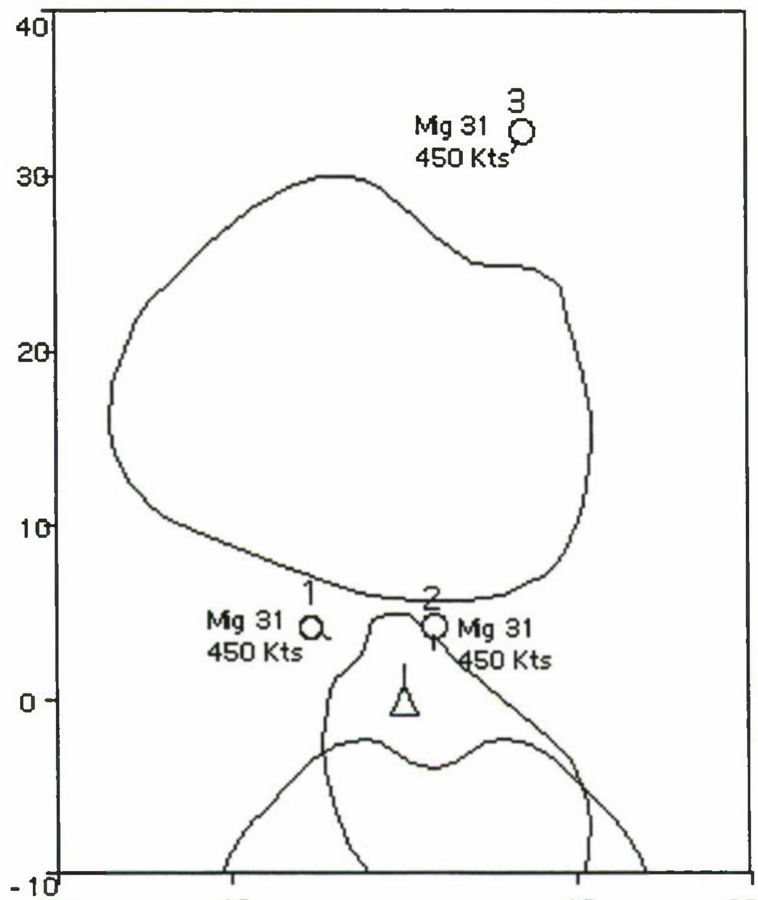


Figure 1. Launch Success Zone (LSZ) Display

Performance was measured in terms of RMS error from the assigned course and total time spent inside the threats' launch success zones. In addition, pilot SA was measured by the Situation Awareness Global Assessment Technique (SAGAT) (Endsley, 1990a) during 4 of the 8 trials in each condition and by the Situational Awareness Rating Technique (14-D SART) (Selcon & Taylor, 1990; Taylor, 1990) during 4 of the 8 trials in each condition. SAGAT and SART were administered during freezes which occurred at random times during the threat avoidance task. In addition, a set of subjective assessments was required of the subjects following each SART administration. These assessments had subjects rate (on a bipolar scale) their overall SA, the sufficiency of their SA, their confidence level regarding their SA, their performance.

In each condition for each subject, during 2 trials, both SART (plus the subjective questions) and SAGAT were administered. During 2 trials only SART (plus the subjective questions) were administered and during 2 trials only SAGAT was administered. During the remaining 2 trials, no SA assessment was done and only performance measures were collected. The order of these administrations was randomized across subjects and conditions.

Results

Performance

Analysis of Variance (ANOVA) was used to evaluate subjects' performance data. There was no difference between conditions on the RMS flight path error. Subjects' deviation from the prescribed path was not affected by the provision of the threat envelopes. They did spend less time in the threats' launch exposure zones, however, when they were provided with this information, $F(1,11) = 19.962$, $p<.001$. Mean time with the LSZ envelope display was 6.5 seconds as compared to 11.75 seconds without the display.

Situation Awareness

SART - An overall SART score was calculated from the 14 individual SART ratings. The SART rating of SA was found to be significantly higher with the LSZ envelopes than without, $F(1,11) = 12.066$, $p<.01$. In examining the underlying dimensions of SART (understanding, supply of resources, demand on resources), it was found that this result was mainly attributable to differences in subject ratings of understanding. With the LSZ envelope display, subject ratings were higher for understanding ($p<.05$), information quantity ($p<.05$), and information quality ($p<.05$).

SAGAT - The subjects' perception of the situation as recorded via SAGAT was compared to the actual situation at the time of each freeze and an assessment of the accuracy of their SA for each SAGAT query calculated. ANOVAs were performed on each query to examine differences in display conditions on SA.

Although there was a subjective impression of higher SA recorded via SART, the SAGAT results show a mixed picture. In terms of Level 1 SA (perception of basic information), subjects had lower SA regarding the location of threat aircraft, $F(1,81) = 3.136$, $p=.08$, with the LSZ envelopes. Most likely the envelopes had a distraction effect and subjects had less SA regarding exact aircraft position. They did, however, show better SA regarding own heading, $F(1,81) = 5.160$, $p=.026$, and own roll attitude, $F(1,81) = 2.726$, $p=.10$, with the LSZ envelopes. This could be due to lower workload in processing the displays or more likely heading and roll changes were more salient in that the threat envelopes changed considerably with changes in these two variables which occurred when making turns.

In terms of Level 2 SA (understanding of the situation), subjects had a better idea of the imminence of the threats, $F(1,81) = 3.286$, $p=.07$, however, they did not have better SA regarding whether the aircraft could launch at them (probably because they could do this fairly well even without the envelopes based on aircraft aspect angle) or of the highest priority threat. (Non-significant trends were in the direction supporting the utility of the LSZ envelopes, however.)

Subjects also showed lower SA with the envelopes on the only Level 3 question asked "which aircraft would be a threat if you stayed on your current course?" Subjects were less able to report this information correctly with the envelopes than without, $F(1,81) = 3.641$, $p=.06$. This could reflect less tendency to project ahead with the envelopes (due to the ability to rely more on the display for that information). It might also reflect a tendency for subjects to be more conservative in judging threats without the envelopes (i.e., perhaps they were more likely to be over complacent with the envelopes).

Comparison of SA Measures - The SART scale was highly correlated ($r^2 = .67$ to $.74$) with a simple subjective SA rating (on a 1-10 scale), an evaluation of the sufficiency of one's SA, and a subjective rating of confidence level. All of these were also highly intercorrelated ($r^2 = .74$ to $.79$). Of the SART components, Mean Understanding was most highly correlated with these factors ($r^2 = .67$ to $.78$). Whatever subjective impression is being tapped by these scales, they appear to draw upon much the same factor. They did vary somewhat, however, in that they were not perfectly correlated. Subjective performance was highly predicted by the subjective SA and SA sufficiency scales ($r^2 = .61$) and less so by a combination of SART and confidence level ($r^2 = .46$).

Using the 48 trials across subjects and conditions in which SART and SAGAT were collected together, a direct analysis was made of how the SART and SAGAT scores compared. First, a component analysis and correlation analysis show that the 13 SAGAT variables collected are fairly independent, agreeing with previous such analyses (Endsley, 1990b) . This means that trying to compile SA queries on different situational aspects into one combined SA variable is not supported. (In support of this, a simple SA score added across queries proved to be of no significance in the envelope vs no envelope comparison).

Each SAGAT variable was therefore treated independently in comparson with the SART score. A regression of SAGAT variables on SART was not significant on any component . There was no relationship between the subjective SART rating and any of the SAGAT variables. Examining the SART components, Mean Understanding, Mean Supply and Mean Demand, again there was no correlation with the SAGAT measures. The subjective assessment of SA derived via SART does not appear to be related to the objective measures of SA provided by SAGAT.

Discussion

This study supports the utility of providing displays that enhance pilot's SA needs. Not only can such displays be easier to create technologically, but can also result in improved performance without encountering the difficulties that combined human/decision support systems encounter. Whereas a calculated "best path" route display may induce significant disconnects in joint/system decision making, the direct display of the LSZ envelopes enhanced pilot decision making by supporting SA. This approach should also be far more robust in supporting dynamic decision making in a complex and changing environment, in addition to being less likely to slow down decision making or result in system induced decision biases.

This study also supports the utility of using a test-battery approach for evaluating display concepts. Simply showing reduced time in the threat zones does not provide needed information on why that was the case or on potential downsides associated with the display concept. SART proved to be of use in predicting this performance benefit. In this case, the benefit of the display was found to be associated primarily with improvements in subjective understanding ratings, as opposed to reduced workload ratings (supply & demand of attention). This finding is useful in that SART could be used in actual flight operations to evaluate design concepts when detailed performance measures are not available.

While subjects rated their SA as higher with the LSZ envelopes, an analysis of the SAGAT data showed that in actuality, while SA may have been higher on some aspects of the situation, it actually was lower in several areas. Subjects had better SA regarding the imminence of the threats and ownship heading and roll. They had poorer SA regarding aircraft location and projecting which aircraft would be a threat in the future. This most likely reflects changes in attention allocation and processing with the new displays. Examining the effects of a particular display or technology on SA can help illuminate these issues to the designer, who may wish to employ alternative design concepts or make modifications to deal with any potential SA tradeoffs. Such SA tradeoffs have been noted in previous studies (Endsley, 1995a) . For instance, in this display a modification to make the threat symbol more salient (e.g., through color coding or increased intensity) in order to improve SA of aircraft location may be recommended from these results. In addition, the effects of displays which directly present high level information (such as the LSZ) on the tendency of pilots to rely on such "present" information to the exclusion of making their own future projections of aircraft states should be further investigated. This effect may be similar to automation complacency effects which have been noted in many other studies (Parasuraman, Molloy, & Singh, 1993) .

By revealing some of these hidden tradeoffs, the SAGAT measure provides a much greater degree of sensitivity and diagnosticity regarding the effects of the display on pilot SA (and potentially performance) than was available from performance measures alone in the limited simulation testing that is normally conducted in such a study. That is, it reveals SA effects that may be important in the long-run for complex mission performance. It should be noted, however, that because SAGAT scoring is based on binomial data (correct or incorrect), more data is needed to reach a level of statistical significance than might be required with other measures. SAGAT is also not currently appropriate for use during actual flight.

This is one of the first studies to directly compare subjective and objective measures of SA, and it is quite interesting that there was in fact no correlation between these measures. The fact that SART, a subjective measure of SA, was highly correlated with confidence level in one's SA and subjective performance has interesting implications for SA measurement. It has been previously suggested that this might be the case (Endsley, 1995a) . This does not mean that subjective SA measures are not useful, however. Such subjective assessments may provide a critical link between SA and performance. That is, a person's subjective assessment of the quality of their SA may be important in determining how a person will choose to act on that SA (conservatively if it is low or boldly if it is high), independent of the actual quality of that SA.

This has been discussed by Christ et. al. (1994) and Endsley and Jones (1997) . As shown in Figure 2, if SA is good and confidence in that SA is high, a person will act to most likely achieve a good outcome (as it will have been possible to make good decisions and plans based on that SA). If with equally good SA the person has a low level of confidence in that SA, however, they most likely will not act on it (choosing to gather more information or behave protectively), and thus be ineffectual. The person with poor SA, if they recognize that it is poor (low confidence level), will correctly choose not to act (or act protectively) and will continue to gather more information to improve SA, thus averting what could be a very bad outcome. The worst situation is that of the person who has poor SA, but has a high level of confidence in that erroneous picture. Not only will this person be likely to act boldly and incorrectly, but often will draw in others who will be fooled by the false confidence. Of the four possible situations, this is the most dangerous. A critical issue, therefore, is ensuring that operators have as good a picture of the situation as possible and that they are able to attribute the correct amount of confidence or certainty to that picture. This is an important issue which has received little attention to date.

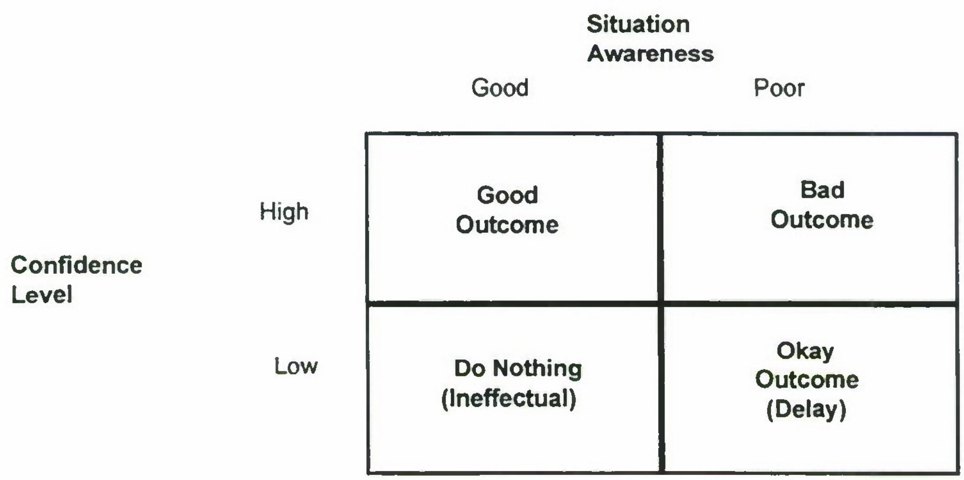


Figure 2. Relationship Between Situation Awareness and Confidence (Endsley & Jones, 1997)

Conclusions

The viability of SA enhancing displays for supporting improved SA, decision making, and performance in the tactical environment was demonstrated by this study. Future research should be directed at creating SA enhancing displays for supporting pilot decision making and for more directly contrasting this approach with traditional expert system approaches. In addition, interesting differences in objective measures of SA and pilot's subjectively reported SA were revealed. These differences may be in themselves an important determinant of behavior in complex tactical scenarios and should be further investigated.

Achnowlegements

This resesarch was supported by a grant from NATO-AGARD.

References

Christ, R. E., McKeever, K. J., & Huff, J. W. (1994). Collective training of multiple team organizational units: Influence of intra- and inter-team processes. In G. E. Bradley & H. W. Hendrick (Eds.), Human Factors in Organizational Design and Management - IV (pp. 323-326). Amsterdam: Elsevier.

Endsley, M. R. (1990a). A methodology for the objective measurement of situation awareness. In Situational Awareness in Aerospace Operations (AGARD-CP-478) (pp. 1/1 - 1/9). Neuilly Sur Seine, France: NATO - AGARD.

Endsley, M. R. (1990b). Situation awareness in dynamic human decision making: Theory and measurement. Unpublished doctoral dissertation, University of Southern California, Los Angeles, CA.

Endsley, M. R. (1995a). Measurement of situation awareness in dynamic systems. Human Factors, 37(1), 65-84.

Endsley, M. R. (1995b). A taxonomy of situation awareness errors. In R. Fuller, N. Johnston, & N. McDonald (Eds.), Human Factors in Aviation Operations (pp. 287-292). Aldershot, England: Avebury Aviation, Ashgate Publishing Ltd.

Endsley, M. R., & Jones, W. M. (1997). Situation awareness, information dominance, and information warfare (Tech Report 97-01). Belmont, MA:

Endsley, M. R., & Kaber, D. K. (1996). Level of automation effects on performance, situation awareness and workload in a dynamic control task. Lubbock, TX: Texas Tech University.

Endsley, M. R., & Kiris, E. O. (1995). The out-of-the-loop performance problem and level of control in automation. Human Factors, 37(2), 381-394.

Fletcher, G. C., Shanks, C. R., & Selcon, S. J. (1996). The validation of an explanatory tool for data-fused displays for high-technology aircraft. In Proceedings of the Head Mounted Displays . Orlando, FL: SPIE.

Hart, S. G., & Sheridan, T. B. (1984). Pilot workload, performance, and aircraft control automation. In Human factors considerations in high performance aircraft (AGARD-CP-371) (pp. 18/1-18/12). Neuilly Sur Seine, France: NATO-AGARD.

Kibbe, M. (1988). Information transfer from intelligent EW displays. In Proceedings of the Human Factors Society 32nd Annual Meeting (pp. 107-110). Santa Monica, CA: Human Factors Society.

Parasuraman, R., Molloy, R., & Singh, I. L. (1993). Performance consequences of automation-induced complacency. International Journal of Aviation Psychology, 3(1), 1-23.

Selcon, S. J. (1990). Decision support in the cockpit: Probably a good thing? In Proceedings of the Human Factors Society 34th Annual Meeting (pp. 46-50). Santa Monica, CA: Human Factors Society.

Selcon, S. J., Bunting, A. J., Coxell, A. W., Lal, R., & Dudfield, H. J. (1995a). Explaining decision support: An experimental evaluation of an explanatory tool for data-fused displays. In Proceedings of the Eighth International Symposium on Aviation Psychology (pp. 92-97). Columbus, OH: The Ohio State University, Department of Aviation.

Selcon, S. J., Smith, F. J., Bunting, A. J., Irving, M. A., & Coxell, A. W. (1995b). An explanatory tool for decision support on data-fused panoramic displays. In Proceedings of the SPIE International Symposium on Aerospace/Defence Sensing and Control and Dual-use Photonics (pp. 17-21). Orlando, FL: SPIE.

Selcon, S. J., & Taylor, R. M. (1990). Evaluation of the situational awareness rating technique (SART) as a tool for aircrew systems design. In Situational Awareness in Aerospace Operations (AGARD-CP-478) (pp. 5/1 -5/8). Neuilly Sur Seine, France: NATO - AGARD.

Taylor, R. M. (1990). Situational awareness rating technique (SART): The development of a tool for aircrew systems design. In Situational Awareness in Aerospace Operations (AGARD-CP-478) (pp. 3/1 - 3/17). Neuilly Sur Seine, France: NATO - AGARD.

Wiener, E. L. (1988). Cockpit automation. In E. L. Wiener & D. C. Nagel (Eds.), Human Factors in Aviation (pp. 433-461). San Diego: Academic Press.

Wiener, E. L. (1993). Life in the second decade of the glass cockpit. In R. S. Jensen & D. Neumeister (Ed.), Proceedings of the Seventh International Symposium on Aviation Psychology (pp. 1-11). Columbus, OH: Department of Aviation, The Ohio State University.

OLIPSA: ON-LINE INTELLIGENT PROCESSOR FOR SITUATION ASSESSMENT

Sandeep Mulgund, Gerard Rinkus,
Christine Illgen, and Greg Zacharias
Charles River Analytics
Cambridge, MA

John Friskie
Wright Laboratory
Wright-Patterson Air Force Base, OH

ABSTRACT

This paper describes a study that assessed the feasibility of developing a concept prototype for an On-Line Intelligent Processor for Situation Assessment (OLIPSA), to serve as a central processor to manage sensors, drive decision-aids, and adapt pilot/vehicle interfaces in the next-generation military cockpit. The approach integrates several enabling technologies to perform the three essential functions of real-time situation assessment:

- Event detection uses a fuzzy logic (FL) processor and an event rulebase to transform fused sensor data into situationally-relevant semantic variables.
- Current situation assessment is performed using a belief network (BN) model to combine detected events into a holistic "picture" of the current situation, for probabilistic reasoning in the presence of uncertainty.
- Future situation prediction is carried out via case-based reasoning, to project the current situation into the future via experience-based outcome prediction.

The prototype OLIPSA was developed in object-oriented C^{++} and integrated with a flight simulation model on an SGI workstation. OLIPSA's performance was demonstrated initially in the defensive reaction portion of an air-to-ground attack mission, in which a pilot must deal with an attack from threat aircraft. Situation awareness (SA) models were developed to support the pilot's assessment of the threat posed by detected aircraft.

INTRODUCTION

Air combat demands that pilots make dynamic decisions under high uncertainty and high time pressure. Under such conditions, numerous empirical studies (Stiffler, 1988) and pilots' own accounts (Shaw & Baines, 1988; Baker, 1986; Singleton, 1990) indicate that the most critical component of decision-making is situation awareness (SA), obtained via the rapid construction of tactical mental models that best explain the accumulating evidence obtained through continual observation of the environment. Once a mental "picture" is developed, decisions are automatically driven by the selection of pre-defined procedures associated with the recognized tactical situation. This is SA-centered decision-making (sometimes called recognition-primed decision-making (RPD)), and it has been widely accepted as the most appropriate representation of actual human decision-making in high tempo, high value situations (Endsley, 1989; Endsley, 1990; 1993;Endsley, 1995b; Fracker, 1990; Klein, 1989a; Klein, 1994; Stiffler, 1988).

As a result, improving pilot SA in air combat has become a principal goal of the Human Systems Technology Investment Recommendations made by USAF's Development Planning Directorate (Development Planning Directorate, 1995). Many new technologies and subsystems are being considered to enhance pilot SA. The problem is not one of a lack of subsystem development efforts; rather it is the lack of an integrated approach to delivering this information to the pilot in a more holistic fashion, and in a manner that reflects the current air combat situation. What is missing is a simplified approach to assessing the tactical situation, so that sensor management, decision-aiding, and pilot/vehicle interface (PVI) management all occur within the overall context set by the assessed situation. What is called for is a computationally intelligent design for an on-line *situation assessment processor*, driving the high-level coordination of all other on-board assets.

The premise of this research is that a situation assessment processor can be developed best by first understanding how the adept *human* pilot accomplishes on-line situation assessment, developing a model of that behavior, and then implementing a version of that model using modern computational intelligence

technologies. Much of the groundwork in the first two areas, understanding and modeling pilot SA, has been conducted over the last decade, and has been extensively reported on in engagement studies (Stiffler, 1988; Hamilton, Dorchak, & Stuart *et al.*, 1988), in pilot self-assessment studies (Baker, 1986; Shaw & Baines, 1988; Singleton, 1990), in behavioral studies of the tactical area (Klein, 1989a; Klein, 1989b; Fracker, 1990; Endsley, 1989; 1990; 1993; 1995); and in recent modeling studies (Zacharias, Miao, Riley & Osgood, 1992; Zacharias *et al.*, 1996; Adams, Tenney & Pew, 1995). This research facilitates implementation of a *pilot-centered* SA processor that synergistically combines knowledge of how pilots accomplish this task with the new-found ability to host intelligent computing technologies in the modern day cockpit.

Development of an on-line situation assessor for the advanced cockpit calls for an integration of several enabling technologies, implementing the three essential functions performed during real-time situation assessment: event detection, current situation assessment, and future situation prediction. The OLIPSA prototype uses fuzzy logic (FL) to implement a "front-end" event detection module, which transforms fused sensor data into situationally-relevant semantic variables (the essential events that, as a group, define the overall tactical situation). OLIPSA uses belief networks (BNs) to implement a current situation assessment module, which combines the detected events with one or more structural models of the tactical situation, to provide a probabilistic assessment of the situation. Finally, OLIPSA uses case-based reasoning (CBR) to project the current assessed situation into the future, over a tactical time window if interest, to support current assessment of future situations, and subsequent plan development and evaluation. In developing a concept prototype to assess overall feasibility of the approach, an object-oriented design was used for software module specification.

BACKGROUND

The development of an on-line situation assessor may be considered a partial solution to the overall **data fusion** problem. The objective of data fusion is to combine data from multiple sources intelligently, to develop a meaningful perception of the environment (Waltz & Llinas, 1990). While humans have long been able to fuse remotely sensed data using mental reasoning methods and manual aids, in recent years there has been considerable interest in developing automated systems capable of combining data from multiple sensors to derive meaningful information not available from any single sensor. This interest has been motivated by the following concerns:

- Increases in target mobility and weapon lethality demand shorter detection/identification and reaction times
- More complex threats demand improved detection/discrimination capabilities
- Increased personnel costs in some missions have dictated remotely controlled or autonomous weapon systems that require data fusion

These demanding requirements and the increasing complexity of available sensor data exceed the human ability to associate and classify incoming data without decision aids, motivating the automation of various data fusion processes. The Joint Directors of Laboratories (JDL) Data Fusion Subpanel have identified three levels of fusion processing products (Waltz & Llinas, 1990; White & Cohen, 1980):

- **Level 1 DF**: Fused position and identity estimates
- **Level 2 DF**: Friendly or hostile military situation assessments
- **Level 3 DF**: Hostile force threat assessments

Across these levels of *information products*, the generality of the results increases from the very specific (e.g., "missile type A at coordinates B") to the more general (e.g., "missile attack in progress on city C"). At level 1, numeric procedures such as estimation or pattern recognition dominate the processing operations. Level 1 information products arise from single and multisource processing (such as target tracking) by sampling the external environment with available sensors and other information sources. The products of this processing are position and identity estimates for targets or platforms in the composite field of view (Waltz & Llinas, 1990). Symbolic reasoning processes involving higher levels of *abstraction* and *inference* dominate the level 2 and 3 fusion operations. Situation abstraction is the construction of a generalized situation representation from incomplete data sets to yield a contextual *interpretation* of level 1 products. This level of inference is concerned with deriving knowledge from some type of pattern analysis of level 1 data. The distinction between levels 2 and 3 is that level 3 products attempt to quantify the

threat's capability and predict its *intent* by projecting into the future, whereas level 2 results seek to indicate *current* hostile behavior patterns.

By reviewing conventional definitions of **situation assessment**, it becomes clear that the level 2 and 3 data fusion functions strongly overlap key situation assessment activities. These include (Noble, 1989): 1) an *estimate* of the purpose of activities in the observed situation, 2) an *understanding* of the roles of the participants in the activities, 3) *inferences* about completed or ongoing activities that cannot be observed directly, and 4) inferences about *future* activities. Endsley, 1987Endsley (1995) has also proposed a general definition of SA:

> *Situation awareness is the perception of the elements in the environment within a volume of time and space, the comprehension of their meaning, and the projection of their status in the near future.*

Each of the three hierarchical phases and primary components of this definition can be made more specific: (Endsley, 1995a):

- **Level 1 SA:** Perception of the elements in the environment – This is the identification of the key elements of "events" that, in combination serve to define the situation. This level serves to semantically tag key situational elements for higher levels of abstraction in subsequent processing.
- **Level 2 SA:** Comprehension of the current situation – This is the combination of the level 1 events into a comprehensive holistic pattern, or tactical situation. This level serves to define the current status in operationally-relevant terms, to support rapid decision-making and action.
- **Level 3 SA:** Projection of future status: projection of the current situation into the future, in an attempt to predict the evolution of the tactical situation. This supports short term planning and option evaluation, when the time permits.

A direct comparison of these three levels of SA with the three levels of DF show that the two functions are clearly distinct at level 1, since DF level 1 focuses on the *numeric* processing of tactical elements to subserve identification and tracking, whereas SA level 1 focuses on the *symbolic* processing of these entities, to identify key "events" in the current situation. At level 2 the definitions are virtually identical, concluding with the conventional definition of SA, that of generating a holistic pattern of the *current* situation. At level 3, the SA definition is more general than the DF definition, since the former also includes projection of ownship and friendly intent, whereas the latter only focuses on threat intent. The following integration of the two DF and SA hierarchies are thus proposed, incorporating a single initial DF level to generate estimates of the tactical elements, and three follow on SA levels, to generate key events, current status, and projected future:

- **Level 1 DF:** Fuse position and identity estimates
- **Level 1 SA:** Perceive elements and identify events
- **Level 2 SA:** Assess (comprehend) current situation
- **Level 3 SA:** Project (predict) future situation

This four-level process progressively raises the level of abstraction of the sensor data and pilot knowledge base, to provide a high level holistic view of the tactical situation.

Figure 1 illustrates how the data fusion and situation assessment functions relate to the other key components of an on-board system. We begin with a specification of the external environment. This includes the specification of all friendly, hostile, and neutral forces, as well as a description of the local terrain and current weather conditions. Information on the external environment is sensed by an on-board sensor suite. The Data Fusion Processes fuse this sensor data to generate individual target tracks and to classify and characterize targets. The Situation Assessment Processor uses this fused track data to generate a current situational state from detected events, and then projects the situation to predict future situations. The total situation assessment (events, current situation, predicted situation) feeds higher-level processing for sensor management, threat assessment, decision-aiding, and pilot/vehicle interface (PVI) adaptation. The dotted lines indicate OLIPSA's envisioned scope.

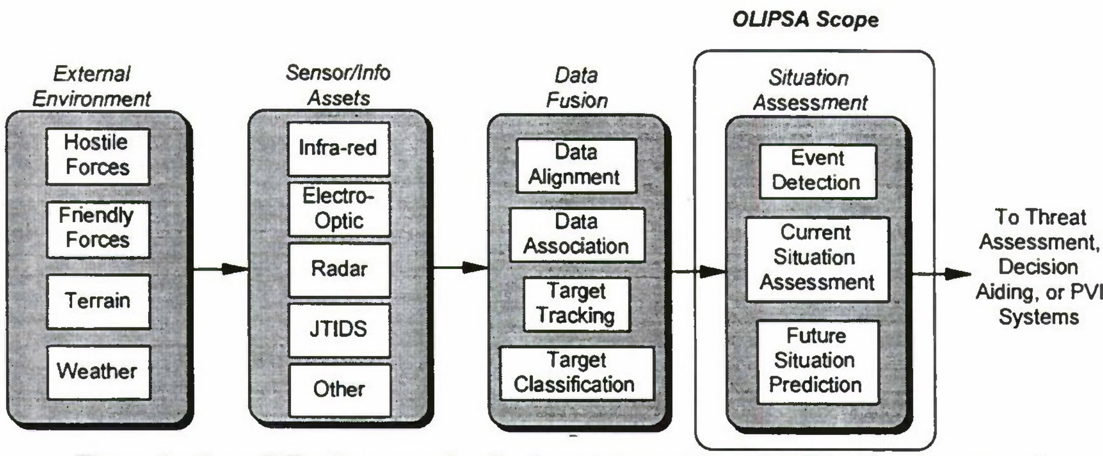


Figure 1. Overall Environment for On-Board Data Fusion and Situation Assessment

OLIPSA FUNCTIONAL DESIGN

Figure 2 illustrates OLIPSA's functional block diagram, which employs a four-stage processor architecture: 1) an event detector, 2) a current situation assessor, 3) a future situation projector, and 4) a context processor. These are described briefly in the following paragraphs.

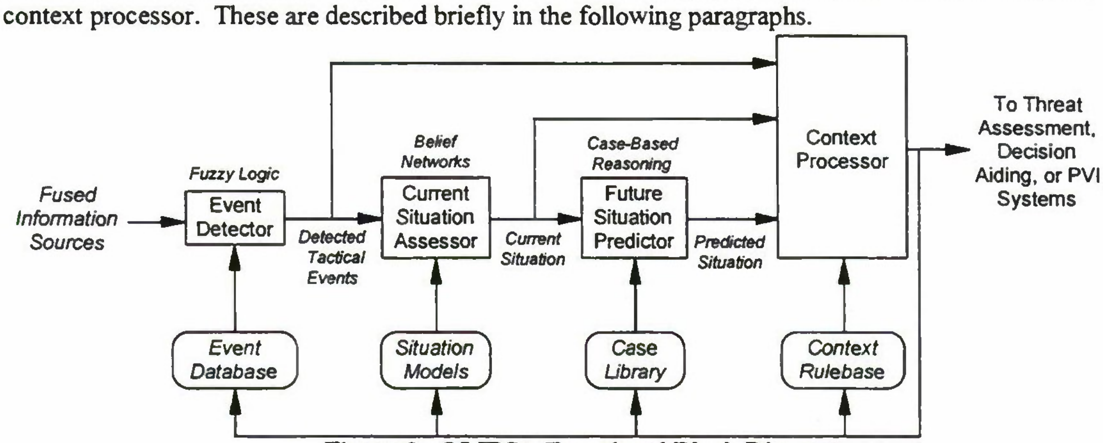


Figure 2. OLIPSA Functional Block Diagram

The **event detector** translates the primarily numerical data generated by the data fusion processor into symbolic data defining key tactical elements. Examples of events are a detection of a threat CAP breaking orbit, a bandit engaging, a mission-related milestone (e.g., crossing the FEBA), or other similar scenario-relevant events that contribute to defining the overall situation. The event detector "engine" can be as simple as a binary threshold logic, converting a numerical value (e.g., threat range) into a Boolean event (e.g., within range of threat envelope). OLIPSA uses a more robust approach to event detection via fuzzy logic (FL) and graded membership functions. Note also the presence of an event database in the figure, which the detector uses to transform scenario states into situation events.

The **current situation assessor** takes in the detected events and generates an assessed situation state S(t), which is a multi-dimensional vector defining the belief values of a number of possible situations currently facing the pilot. The situations, their relation to one another, and their association with detected events are all defined by a set of situation models, each model being a tree of possible situations and events. OLIPSA employs belief networks (BNs) to implement both the situation models and the situation assessment function. This provides a way of making computationally explicit the extremely complex process of situation assessment in a real-time environment, while at the same time ensuring a fair degree of rigor in inferencing via the use of Bayesian reasoning logic. The net result of this stage of processing is the generation of an aggregated set of situation likelihoods (belief values) and their associated event probabilities, which define the overall situation.

The **future situation predictor** takes in the assessed current situation and generates, over a limited time window, a prediction of the most likely future situations, or $\{\mathbf{S}(t+T), \mathbf{S}(t+2T), \ldots \mathbf{S}(t+NT)\}$. The output of this block is at the same level of granularity as the proceeding block (i.e., at the situation level rather than the multi-vehicle state level, for instance), but the output future situations are also indexed against future time, from the current time to the end of the limited time window. This provides for a view of how the current tactical situation can be expected to evolve over the time window of interest. The conventional engineering approach for this type of predictor would couple an engagement level simulator with a *current* situation assessor module, and generate, via multiple simulation runs, likely scenarios spanning out into the future. Unfortunately, this is computationally infeasible, given the complexity of engagement level models, and the requirement for real-time on-board operation. OLIPSA employs a computationally less intensive approach via the use of case-based reasoning (CBR), and a well-stocked case library of possible engagement scenarios. This allows OLIPSA to rapidly index the current situation, and on the basis of comparable situations maintained in the library, generate a likely evolution of the current situation, using simple transition rules, at a high level of granularity defined by the situational parameters. The net result is an *experience-based prediction* into the future, with a minimum of computational overhead.

Lastly, the diagram shows a **context processor**, which takes in the outputs of the preceding blocks and generates, via a conventional production rule system, an overall assessment of the situation. This last stage of processing allows the introduction of additional heuristics not accounted for in the upstream processing blocks, provides a means of resolving conflicts in the different blocks, and, perhaps most importantly, provides a path by which feedback can be given to the three upstream knowledge bases: the event database, the situation models, and the case library. This feedback path allows for the eventual incorporation of learning and adaptation techniques.

Event Detection using Fuzzy Logic

OLIPSA uses fuzzy logic (FL) technology to implement the event detector module. While some event detection, such as subsystem discretes (e.g., ownship missile launch), can be implemented via simple discrete Boolean logic, most tactically-significant events require a more robust and flexible means of expression. This can be achieved by the use of fuzzy logic.

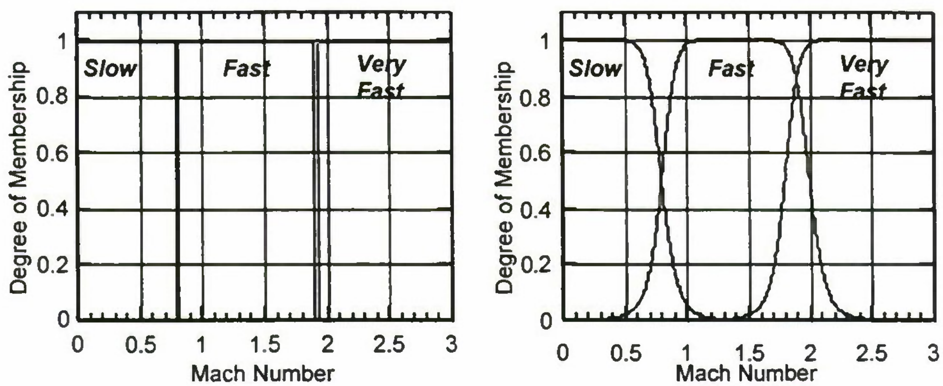
Fuzzy logic was proposed by Zadeh (1965, 1973) as a mathematical concept to deal with uncertainty in human decision-making. He was concerned with how humans process imprecise non-numerical, or linguistic, information (i.e., *big*, *small*, *very fast*, *heavy*, etc.) to perform a given task. He argued that if a human can perform complex tasks with this imprecise knowledge, then a machine would also benefit from such an approach. Zadeh defined multi-valued or fuzzy sets that are defined by a *membership function*. This membership function relates to or measures the *degree* to which a given element belongs to a set (unlike conventional sets, which specify only whether or not a particular element belongs to some set). Later, he introduced the concept of linguistic variables. For instance the variable pitch attitude could take values of *negative big*, *negative small*, *zero*, *positive small*, *positive big*, etc. Fuzzy logic has been successfully applied in such areas as statistical analysis, pattern recognition, image analysis, robotics, and control theory.

The OLIPSA prototype uses fuzzy logic for deriving the low-level data that feeds the threat assessment network (described below). In the BN formalism, continuous readings must be quantized into one of a finite set of descriptions; for example, speed is characterized as *slow* (Mach 0 to 0.8), *fast* (Mach 0.8 to 1.9), and *very fast* (Mach 1.9 and beyond). However, it is intuitively apparent that there is no meaningful difference between a speed of Mach 0.799 and Mach 0.801. Unfortunately, a BN that uses "hard" boundaries between one category and the next would classify these values as slow and fast, respectively, which might lead to different conclusions regarding vehicle type. Further, if a contact is accelerating or decelerating smoothly, a transition from slow to fast would result in discontinuous jumps in BN output, which may produce undesirable effects.

For example, Figure 3a illustrates the speed classifications using classical, non-fuzzy sets. A given speed measurement may fall into only one of the three categories shown. By contrast, figure 5b shows three fuzzy sets for the classifications *slow*, *fast*, and *very fast*. A given speed measurement is characterized by its degree of membership in each of these fuzzy sets. For example, a speed of Mach 0.8

has a 0.5 degree of membership in the fuzzy set *slow*, and a 0.5 degree of membership in the fuzzy set *fast*. As speed increases or decreases, these degrees of membership change in a continuous manner. This fuzzy characterization has the following benefits:

- The degrees of fuzzy membership are tolerant of noise in the speed measurement.
- Similar inputs produce similar outputs, to maintain a smooth, non-jerky response from the BN to time-varying input signals.
- It facilitates knowledge engineering, as it mimics the human approximation process.



(a) Non-Fuzzy Classification (b) Fuzzy Classification

Figure 3. Classification of Threat Speed

Current Situation Assessment using Belief Networks

The second key component of OLIPSA's architecture is the current situation assessment module, which uses aircraft information system outputs to generate a high-level interpretation of the tactical situation facing the pilot. The OLIPSA prototype relies on belief networks (Pearl, 1988) for reasoning in the presence of uncertainty.

Any robust computational model of situation assessment requires a technology that has: 1) a capability to quantitatively represent the key SA concepts such as situations, events, and the pilot's *mental model*; 2) a mechanism to reflect both diagnostic and inferential reasoning; and 3) an ability to deal with various levels and types of uncertainties, since most real-world systems of any complexity involve uncertainty. Russell & Norvig (1995) cite three principal reasons for this uncertainty:

- *Theoretical ignorance*: All models of physical systems are necessarily approximations.
- *Laziness*: Truly exceptionless rules require numerous antecedents and consequents [cf. 'Frame Problem' (McCarthy & Hayes, 1969) and are therefore computationally intractable].
- *Practical ignorance*: Even if all rules are known, we do not always have time to measure all properties of the particular objects that need to be reasoned over.

The principal advantages of belief networks over other uncertain reasoning methods are:

- Its probability estimates are guaranteed to be **consistent with probability theory**.
- It is **computationally tractable**. Its efficiency stems principally from exploitation of conditional independence relationships over the domain.
- The structure of a BN **captures the cause-effect relationships** that exist among the variables of the domain. The ease of causal interpretation in BN models typically makes them easier to construct (Henrion, 1989).
- The BN formalism **supports many reasoning modes**: causal reasoning from causes to effects, diagnostic reasoning from effects to causes, mixed causal and diagnostic reasoning, and intercausal reasoning. Intercausal reasoning refers to the situation in which a model contains two potential causes

for a given effect. No other uncertain reasoning formalism supports this range of reasoning modes (Russell & Norvig, 1995).

Belief networks provide the capability and flexibility of modeling SA with its full richness. They also provide a comprehensible picture of the SA problem by indicating dependent relationships among variables, at both high-levels (symbolic) and low-levels (numeric). This provides a clearer view of how each individual piece of evidence affects the high-level situation characterization. They allow the incremental addition of evidence at any network node as it arrives, thus allowing for real-time SA update. Finally, BNs enable a designer to partition a large knowledge base into small clusters, and then specify probabilistic relationships among variables in each cluster (and between neighboring clusters). This approach facilitates construction of large, robust knowledge bases without explicitly specifying the relationships between all possible combinations of variables. This feature is especially useful for large domains such as tactical situation assessment.

Future Situation Prediction using Case-Based Reasoning

An important component of situation awareness—and a major goal of the OLIPSA system—is the ability to accurately predict, faster than real-time, the evolution of the current situation over some future window of time, in sufficient detail to guide the pilot's decisions. Reliable prediction, even over a time scale of several seconds, could provide a significant tactical advantage to a strike package. Furthermore, it is clear that the longer the prediction time, the greater the advantage. However, this is a formidable prediction problem: the multiple-entity air combat environment is a complex dynamic system whose evolution is determined not only by the low-level dynamics of the entities (e.g., aircraft, missiles, etc.) but also by *intelligent* decision-making processes occurring in the minds of the individual combatants and at higher (tactical) levels involving multiple aircraft.

In the development of the OLIPSA prototype, we evaluated the use of case-based reasoning (CBR) for situation prediction. CBR allows us to index the current situation against a library of past tactical scenarios, and on the basis of comparable situations maintained in the library, generate a likely evolution of the current situation, using simple transition rules, at a high level of granularity defined by the situational parameters. The net result is an *experience-based projection* into the future, with a minimum of computational overhead, thus making the approach particularly amenable to real-time implementation with limited computational resources.

A fundamental principle of prediction motivating our approach can be seen in figures 4a and b. Shown are two prediction periods for a given snapshot of a tactical situation. Very simply, the smaller the prediction time, the less important are the inter-entity interactions/relationships that carry a great deal of information about the enemy's tactics and plan. Over very short intervals, the physics of the situation dominates its evolution and individual aircraft trajectories can be computed, to close approximation, independently of each other. Figure 4a depicts this situation: the short gray arrows indicate a small prediction time. As the prediction time increases, the inter-entity information (e.g., the intercept tactic the enemy aircraft are using) becomes the most important factor in predicting the evolving situation. This is suggested in figure 4b, where Dt is larger. Knowledge of the maneuver the enemy is currently employing greatly increases the accuracy and confidence in the predicted state of the system at larger Dt. Thus, to maximize the prediction time and the associated tactical advantage, we should include as much high level information describing inter-entity relationships as possible in the model. The OLIPSA approach is also founded on the following assumptions:

- The air combat situation space can be represented fairly coarsely and still provide highly useful—i.e., mission critical—information on which each combatant can act. This assumption is mandated by the need for real-time performance. If the behavior of all aircraft, missiles, etc., in the scenario must be characterized in great detail, faster-than-real-time predictions will be infeasible. Thus, a fairly coarse discretization of the airspace and of time is necessary.
- Air combat is constrained by tactical principles, in the sense that military pilots undergo extensive training in which the accumulated wisdom and experience are inculcated in them. Thus, we can expect that past scenarios will be good predictors of new engagements.
- The longer the prediction time, the better the plans the pilot can generate.

- Assuming that the threat aircraft are executing some formation tactics, it follows that the movements/actions of any one aircraft will generally be predictive of the movements of the others. In other words, there is a great deal more information—and thus, more predictive power—in the spatio-temporal correlations over the *set* of enemy aircraft than in the individual trajectories viewed in isolation. Dong & Atick (1995) provide analyses showing this is generally true of spatio-temporal image sequences. Accordingly, any proposed prediction system should efficiently represent spatio-temporal context.

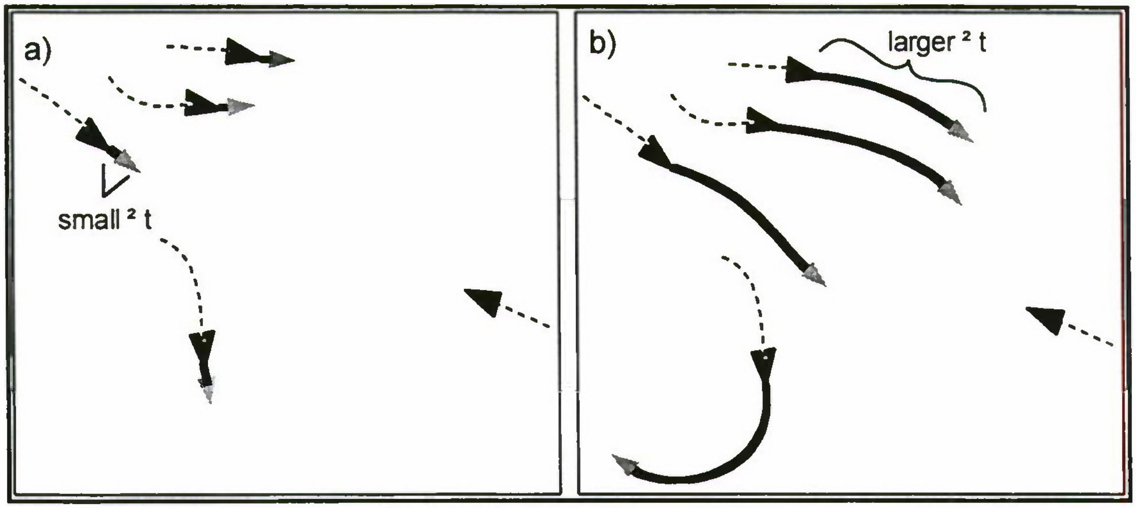


Figure 4. Trajectory Prediction over (a) Short and (b) Long Prediction Time

OLIPSA IMPLEMENTATION FOR AIR-TO-AIR THREAT ASSESSMENT

The prototype OLIPSA models the threat assessment process that a fighter pilot carries out during the defensive reaction segment of an air-to-ground attack scenario. During this segment of the scenario, the pilot must make a judgment as to the potential threat to ownship posed by an aircraft detected by the onboard sensors, in the overall context of the mission and its rules of engagement. This assessment will support the pilot's decision to attack, avoid, or defend against the detected contact.

Figure 5 illustrates the threat assessment BN we developed through consultation with a subject matter expert. The network quantizes threat potential into one of four categories: *high*, *medium*, *low*, or *none*. The *rectangular* nodes denote information that is derived directly from sensor measurements, while the *oval* nodes denote hypotheses that are computed in accordance with the axioms of probability from this sensor data, using the information stored in the network's conditional probability tables.

The threat potential depends directly on the type of sensor contact (*transport*, *bomber*, *fighter*, or *missile*), its aspect angle (*low*, *medium*, or *high*), its location with respect to our estimate of its threat envelope (*out of range*, *within range*, *inside minimum range*, or *within gun range*), and its maneuvering actions (*turning towards ownship*, *turning away*, or *non-maneuvering*). The *type* and *aspect* hypotheses are integrated into a single super-node called *type and aspect* because fighter aircraft and missiles are a greater threat when pointing directly at ownship, while bombers often have tail guns and, therefore, are of greater concern when pointed directly away. As such, the effect of aspect angle on threat hypothesis depends on the type of vehicle, making it appropriate to integrate the two variables. The *type* hypothesis depends on the following: 1) the output of ownship's non-cooperative target recognition system (NCTR), 2) the radar type classification generated by the radar warning receiver (RWR), if any, 3) the contact's altitude, and 4) the contact's speed. Contact altitude, speed, and aspect angle are derived using radar system outputs and ownship air-data systems.

In the course of a simulation, multiple copies of this network are instantiated (one for each sensor contact). The threat hypothesis generated by each network drives the radar display symbology, so that display appearance correlates with the hypothesized threat assessment. The intent of this adaptation is to

make display symbology relate to the situation, to assist the pilot in assessing which detected contacts pose the most potential hazard to mission completion.

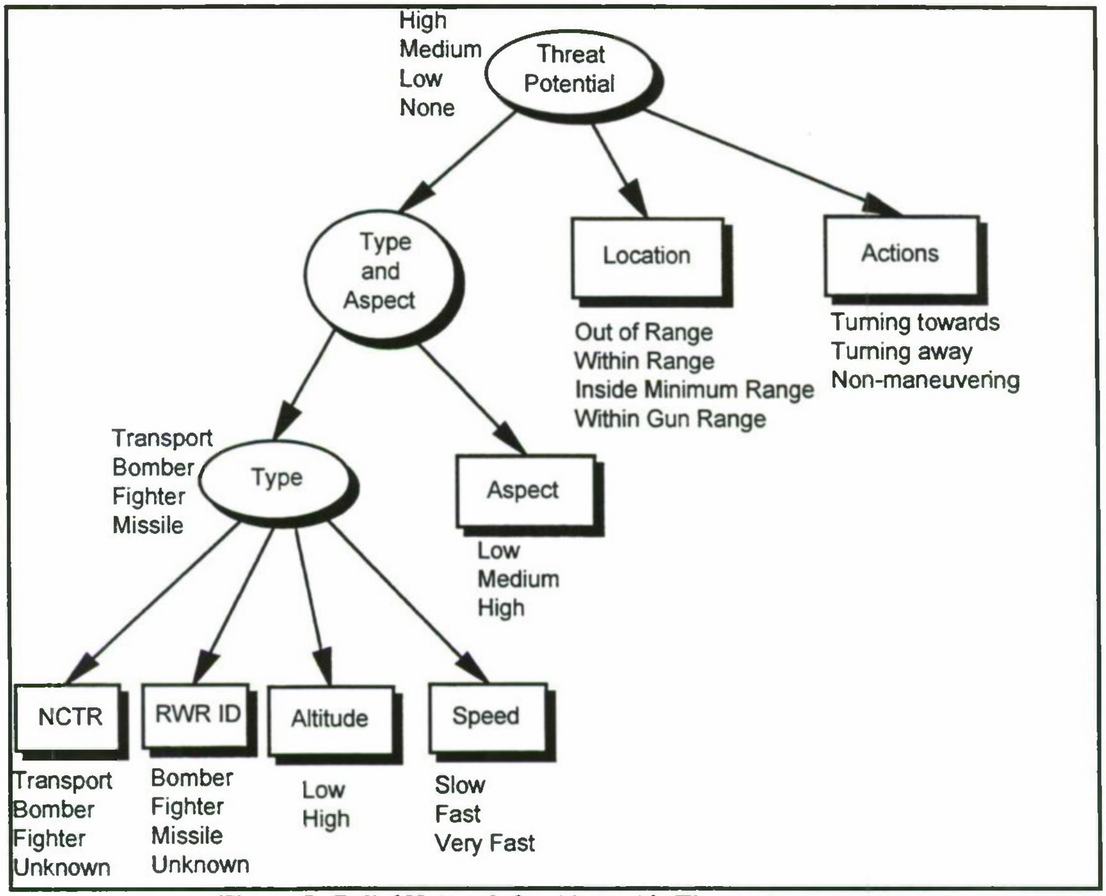


Figure 5. Belief Network for Air-to-Air Threat Assessment

Figure 6 depicts the sample case structure based on the threat assessment BN developed under this effort. These variables encode an approximate description of each enemy entity's position, aspect, etc., relative to the ownship. Therefore, the entire spatio-temporal vector comprising the case implicitly represents all higher, plan- and tactical-level information present in the system. Other high-level features of the battle scenario may also be included in the case description.

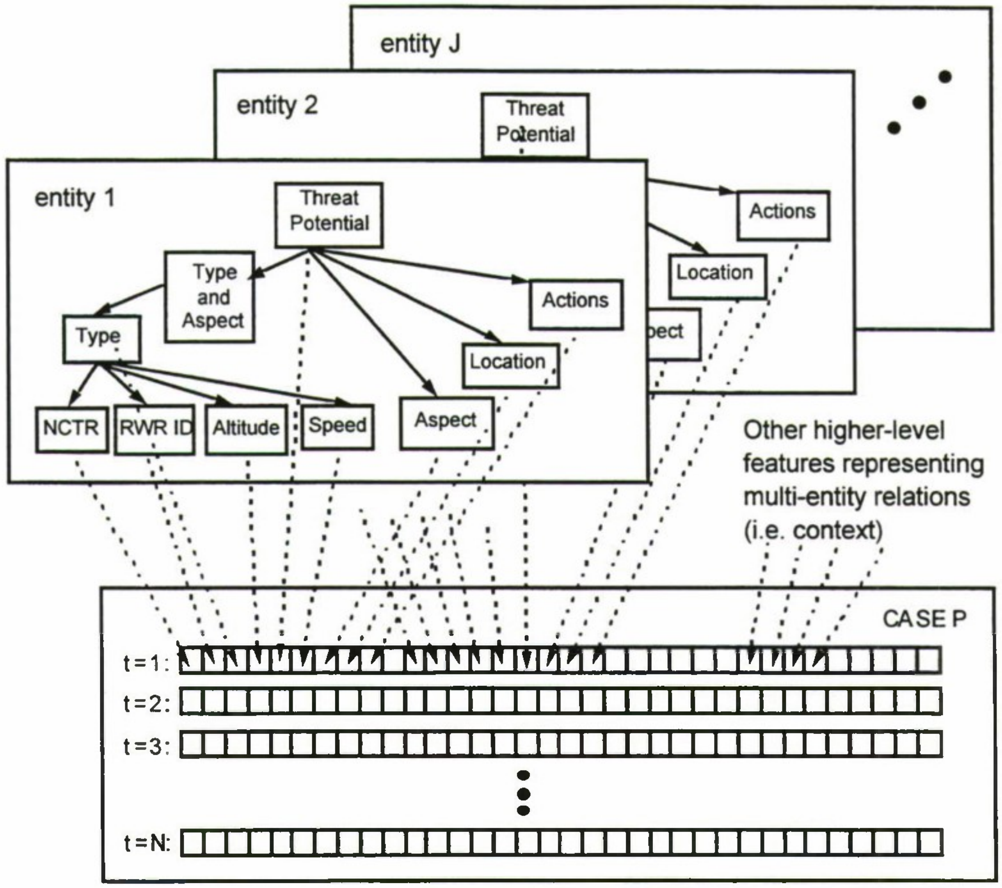


Figure 6. Case Structure For OLIPSA Future Situation Prediction Module

The prediction module's initial development was carried out using the PC-based *Esteem* CBR development environment. This standalone shell allows rapid creation and analysis of a broad range of case definitions, similarity metrics, and adaptation metrics. In particular, it allows us to compare prediction performance across a range of index set sizes.

The desired output from the prediction module is the most likely evolution of the situation over the next temporal window. Accordingly, the general framework is to compare the vector describing the current situation to the situations stored in the library of past cases, returning the case containing the closest match. For purposes of making the match computation faster, the similarity metric may be defined over a subset of the features rather than all of them. This subset is referred to as the set of index features. This implies less flexibility in the matching process; i.e., less tolerance to missing features. Therefore, some approaches, notably Waltz's memory-based reasoning (Waltz, 1989), include all case features in the index set, thus comprising a very high dimensional *nearest neighbors* technique. Once the closest matching time slice has been found, the remaining time slices in that case are read out as the model's prediction of the scenario's evolution.

OLIPSA-Based System Automation

For initial demonstration of OLIPSA operation, we have developed four levels of SA-driven aiding that are implemented in a prototype air-to-air radar display:

Level 1: Basic air-to-air radar display

Level 2: Graphical radar symbology that incorporates the BN-driven threat assessment, so that symbol coding is proportional to threat hypothesis

Level 3: Augmentation of adaptive graphical radar symbology with audio alerts

Level 4: Adaptive graphical radar symbology, audio alerts, and computer-based control of radar scan pattern as a function of the tactical situation

Level 5: Adaptive graphical radar symbology, audio alerts, computer-based control of radar scan pattern as a function of the tactical situation, and computer takes over task of releasing chaff during missile evasion maneuvers

Figure 7 presents a legend for our air-to-air radar symbol coding, for level 2 aiding. This legend is based in part on the symbology used in the FITE simulator facility at Wright-Patterson Air Force Base (Haas *et al.*, 1995). To this we have added the threat potential coding: as the threat potential predicted by the threat assessment BN increases, the symbol line thickness increases in proportion. For a *high* threat, the symbol is drawn completely filled in, to give it the highest possible salience.

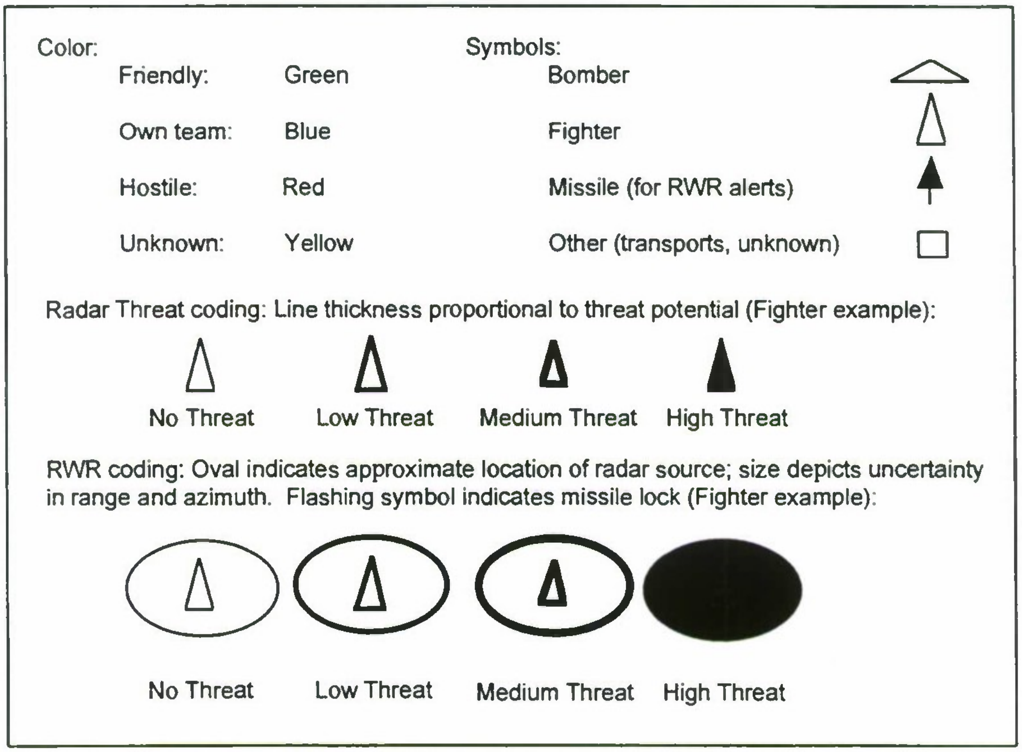


Figure 7. Air-to-Air Radar Symbology

When the RWR detects a radar source, an oval is superimposed on the display to indicate the source's estimated origin. The lower part of the figure shows the RWR oval superimposed on a radar icon. Again, line thickness is proportional to the estimated threat potential. The size of the oval in a direction parallel to the radar line-of-sight indicates uncertainty in range, while the size in the direction perpendicular to the line-of-sight (in the horizontal plane) indicates the uncertainty in azimuth. When the RWR detects a missile lock, the icon will flash, to draw attention to the expected origin of the incoming missile. In the event that the RWR detects a radar signal at a distance beyond the current radar range setting, an arrow appears at the boundary of the radar display to indicate the direction of the incoming radar signal. The arrow's thickness is proportional to the estimated threat potential.

For level 3 aiding, we have incorporated the following audio alerts into our prototype:

- An audio warning tone that sounds when the RWR detects a hostile radar signal at a range beyond the current radar display setting
- An audio alert tone that sounds when the RWR detects an enemy missile lock on ownship

Both audio alerts are correlated with visual icons that indicate the same information. Experimental research has shown that correlated bi-modal alerting provides improved response latencies and enhanced subjective SA over uni-modal alerting (Selcon, Taylor, & Shadrake, 1992). These alerts can be localized using 3-D audio, so that the origin of the sound corresponds to the direction of the detected

threat. There are at least two potential modes of operation for such alerting: 1) bi-modal alerts are presented in all cases, or 2) bi-modal alerts are presented only when pilot workload increases beyond a certain threshold and/or the density of information on the radar beyond a certain level.

For level 4 aiding, the OLIPSA prototype can automatically select appropriate radar display range, in the event that an enemy radar at or above a given threat threshold is detected at a range beyond the current display setting. Again, this automation can happen whenever the tactical situation and system configuration warrants it, or when the pilot workload rises beyond a certain threshold. Finally, for level 5 adaptation, our prototype system can automatically release countermeasures to confuse threat missile sensors in the course of a missile evasion maneuver.

Figure 8 presents a snapshot of our integrated display during a simulation. The italicized annotations show the meaning of the various symbology elements. At the instant shown, an incoming missile is approximately 7.5 nautical miles away from ownship, while the aircraft that fired it is veering to its left. At the same time, the RWR has detected a medium threat fighter radar (as indicated by the "F" within the arrow) at a range greater than 20 nm (the current radar display setting).

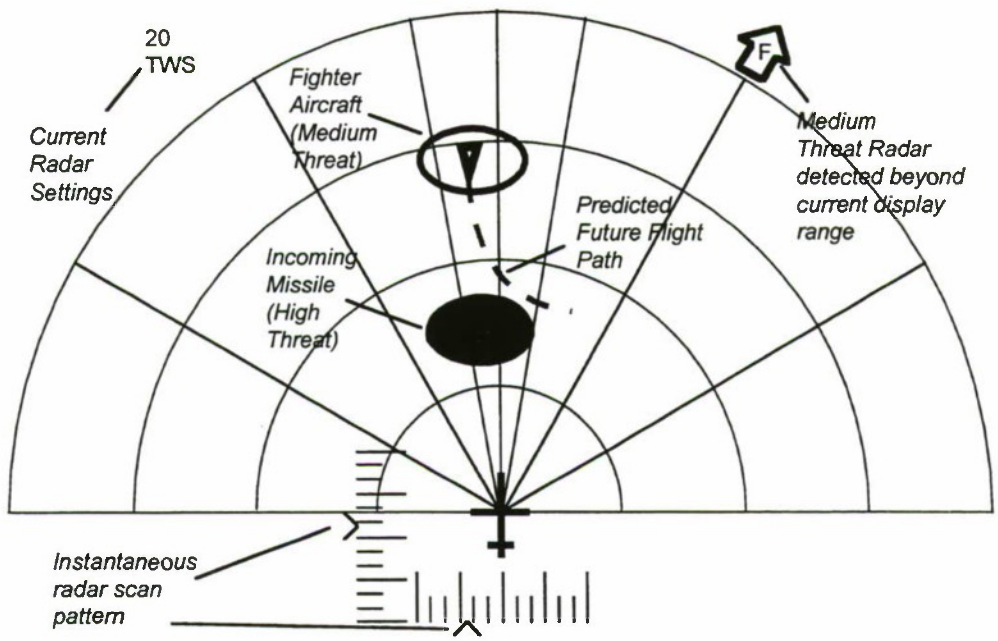


Figure 8. Prototype Radar Display

CONCLUSIONS

This study has demonstrated the basic feasibility of a concept prototype for on-line situation assessment. OLIPSA uses belief networks and fuzzy logic for a robust, extensible framework for current situation assessment modeling. Belief networks readily facilitate extending existing models with new variables or dependencies without re-coding. This offers a considerable benefit over conventional expert system approaches of domain knowledge modeling, in which it is necessary to redesign rulebases whenever a new variable is introduced. The use of fuzzy logic provides an SA module that generates smooth outputs for smooth inputs. Case-based reasoning provides a practical means of implementing experience-based outcome prediction.

The OLIPSA prototype shows how an on-line SA processor can be used for diverse in-cockpit functions such as display content generation, multi-modal alerting, sensor management, and task automation. Each of these functions was demonstrated on a prototype air-to-air radar display. OLIPSA provides a set of conceptual building blocks for follow-on development, which we are now using to develop situation assessment models for uninhabited combat air vehicle (UCAV) operations.

ACKNOWLEDGMENTS

This work was supported by the Wright Laboratory at Wright-Patterson Air Force Base, under Contract F33615-96-C-3607, and by the Armstrong Laboratory under Contract F41624-96-C-6020. The authors are grateful to Professor Mica Endsley for her input on concepts, and to Mr. Robert Shaw for his subject matter expertise on fighter aircraft operations.

REFERENCES

Adams, M. J., Tenney, Y. J., & Pew, R. W. (1995). Situation Awareness and the Cognitive Management of Complex Systems, *Human Factors*, 37(1), 85-104.

Baker, J. R. (1986). Tactics and The Fog of War, *USAF Fighter Weapons Review*, Summer, 14-17.

Development Planning Directorate. (1995). *Technology Investment Recommendations Report*, U.S. Air Force, Wright Patterson Air Force Base.

Dong. D.W. & Atick, J.L. (1995). Statistics of natural time-varying images. *Network: Computation in Neural Systems*, 6(3), 345-358.

Endsley, M. R. (1989). Pilot Situation Awareness: The Challenges for the Training Community, *Interservice/Industry Training Systems Conference*, Ft. Worth, TX.

Endsley, M. R. (1990). Predictive Utility of an Objective Measure of Situational Awareness, *Proc. of the Human Factors Society 34th Annual Meeting*, Santa Monica, CA.

Endsley, M. R., & Bolstand, C. A. (1992). Human Capabilities and Limitations in Situation Awareness, *AGARD: Combat Automation for Airborne Weapons Systems: Man/Machine Interface Trends and Technologies*, Neuilly Sur Siene, France.

Endsley, M. R. (1993). A Survey of Situation Awareness Requirements in Air-to-Air Combat Fighters, *International Journal of Aviation Psychology*, 3, 157-168.

Endsley, M. R. (1995). Toward a Theory of Situation Awareness in Dynamic Systems, *Human Factors*, 35(1), 32-45.

Fracker, M. L. (1990). Attention Gradients in Situation Awareness, *Situational Awareness in Aerospace Operations AGARD-CP-478*, . Neuilly-Sur-Siene, France: Advisory Group for Aerospace Research and Development.

Haas, M. W. (1995). Virtually-Augmented Interfaces for Tactical Aircraft, *Biological Psychology*, 40, 229-238.

Henrion, M. (1989). Some Practical Issues in Constructing Belief Networks, In Kanal, Levitt, & Lemmer (Eds.), *Uncertainty in Artificial Intelligence 3*. North Holland: Elsevier Science Publishers.

Klein, G. A. (1989a). Recognition-Primed Decisions, *Advances in Man-Machine Systems Research*, 5, 47-92.

Klein, G. A., *et al* (1989b). Critical Decision Method for Eliciting Knowledge, *IEEE Trans. on Systems, Man and Cybernetics*, 19(3), 462-472.

Klein, G. (1994). Naturalistic Decision Making: Implications for Design, *Gateway*, 5(1), 6-7.

McCarthy, J., & Hayes, P. J. (1969). Some Philosophical Problems from the Standpoint of Artificial Intelligence, In Meltzer, Mitchie, & Swann (Eds.), *Machine Intelligence 4*, (pp. 463-502). Edinburgh, Scotland: Edinburgh University Press.

Noble, D. (1989). Schema-Based Knowledge Elicitation for Planning and Situation Assessment Aids, *IEEE Trans. on Systems, Man & Cybernetics*, 19(3), 612-622.

Pearl, J. (1988). *Probabilistic Reasoning in Intelligent Systems: Networks of Plausible Inference*. San Mateo, CA: Morgan Kaufmann.

Russell, P., & Norvig, P. (1995). *Artificial Intelligence: A Modern Approach*. Englewood Cliffs, NJ: Prentice Hall.

Shaw, M. L. G., & Baines, B. R. (1988). KITTEN: Knowledge Initiation and Transfer Tools for Experts and Novices, In Boose & Gaines (Eds.), *Knowledge Acquisition Tools for ESs*. NY: Academic Press.

Singleton, J. (1990). The Four Levels of Wingmanship, *USAF Fighter Weapons Review* (Spring).

Stiffler, D. (1988). Graduate Level Situation Awareness, *USAF Fighter Weapons Review* (Summer).

Waltz, E., & Llinas, J. (1990). *Multisensor Data Fusion*. Norwood, MA: Artech House, Inc.

White, L. J., & Cohen, E. I. (1980). A Domain Stretegy for Computer Program Testing, *IEEE Trans on Software Engineering*, SE-6(3).

Zacharias, G., Miao, A., Illgen, C., Yara, J., and Siouris, G. (1996). SAMPLE: Situation Awareness Model for Pilot-in-the-Loop Evaluation, *First Annual Conference on Situation Awareness in the Tactical Air Environment*, Naval Air Warfare Center, Patuxent River, MD.

Zadeh, L. A. (1965). Fuzzy Sets, *Information and Control*, 8, 338-353.

Zadeh, L. A. (1973). Outline of a New Approach to the Analysis of Complex Systems and Decision Processes, *IEEE Transactions on Systems, Man and Cybernetics*, SMC-3(1): 28-44.

DESIGN ISSUES FOR A DECISION SUPPORT SYSTEM FOR A MODERN FRIGATE

Bruce A. Chalmers
Defence Research Establishment Valcartier (DREV)
Decision Support Technologies Section
Val-Belair, Québec, Canada

ABSTRACT

Technological advances in threat technology, the increasing tempo and diversity of open-ocean and littoral scenarios, and the volume and imperfect nature of data to be processed under time-critical conditions pose significant challenges for future shipboard Command and Control Systems (CCSs) and the operators who must use these systems to defend their ship and fulfil their mission. To address these challenges, we are investigating a diverse range of concepts for designing and evaluating a real-time decision support system (DSS), to be integrated into the ship's existing CCS, aimed at providing enhanced decision support capabilities to combat system operators. These capabilities include: i) fusion of data from the ship's sensors and other sources; ii) formulation, maintenance and display of an accurate dynamic situation picture, leading to enhanced situation awareness of operators; iii) identification and selection of courses of action in response to anticipated or actual threats to the mission; and iv) action implementation once a decision to act has been made and is being carried out.

This paper examines a range of issues currently being investigated for the design of the DSS, focusing on automation, cognitive and methodological issues. Automation issues address principles and paradigms for automated decision support. Cognitive issues deal with the specifics of the various cognitive-level behaviours which the DSS must exhibit and/or support. The emphasis is on a cognitively-based model of the Command and Control (C2) process that appears promising for guiding system design. Methodological issues are concerned with managing the complexity of the design problem in a systematic and effective manner. Some key ideas of a specific model-based framework for design are exposed. The results of this research are expected to contribute to DREV's ongoing investigation of enhancements to the Canadian Patrol Frigate's CCS as part of its mid-life upgrade in the next century.

Keywords: command and control process, data fusion, situation awareness, situation representation, action management, decision support, automation, model-based design methodologies, ecological interface design.

INTRODUCTION

Technological advances in threat technology, the increasing tempo and diversity of open-ocean and littoral (i.e., near land) scenarios, and the volume and imperfect nature of data to be processed under time-critical conditions pose significant challenges for future shipboard Command and Control Systems (CCSs) and the combat system operators who must use these systems to defend their ship and fulfil their mission. For the purposes of this paper, the CCS is viewed as a sub-system at the heart of the ship's combat system which includes various other sub-systems like weapon and sensor systems, a navigation system, and an environment monitoring system. The CCS provides automated capabilities to allow operators to use the fighting resources of the ship. However, current operational systems generally provide little support for tactical decision making in complex, highly dynamic scenarios where time for decision making and action execution is at a premium. The need for such support is all the more pressing given the current emphasis on littoral warfare that results in reduced reaction times and the need to deal quickly and correctly with complex rules of engagement designed to avoid undesirable consequences[i].

The Data Fusion and Resource Management Group at Defence Research Establishment Valcartier (DREV), with its industrial and university collaborators, have for several years now been investigating algorithms to augment or enhance existing CCS capabilities, by continuously fusing data from the ship's sensors and other sources, dynamically maintaining a tactical picture, and supporting response to actual or anticipated threats. The emphasis of this work has been on automated capabilities that work in semi-autonomous control mode, with the operator playing a mostly passive, supervisory role. Consequently, operator-in-the-loop issues and their impact on system design have not previously received detailed consideration.

DREV is now broadening the scope of this work. It is involved in a new project to design, develop and implement a real-time decision support system (DSS) that can be integrated into the ship's CCS to assist operators in conducting the tactical Command and Control (C2) process, focusing on Above-Water Warfare (AWW). The operator serves three primary roles in the C2 process: situation interpreter, decision maker and effector. The purpose of the DSS is to support operators in each of these roles. A key goal is to engineer a joint system, comprised of both operators and automated decision aids, that optimises overall mission performance, leading to improved operational effectiveness. Importantly, this work extends the scope of the problem to include human-machine interaction and team-machine collaborative issues, particularly where higher level cognitive processing involving judgements and decision making are involved.

Developing a DSS is extremely challenging for a number of reasons. We examine a few of these reasons here. Reminiscent of Gestalt principles in perception, overall operator-machine performance is an emergent property. Their combined performance emerges from interactions of operators with an external battle environment, with the aid of the DSS, as technology dictates what users can do and as users exploit technological support tools as aids in achieving their mission. The joint operator-machine system is therefore more than the sum of its parts[2]. This suggests that it will be difficult to make effective system design decisions simply on the basis of narrow, localised assessments of performance gains to be derived by special purpose technological solutions. Ignoring this point (as it often is) can lead to problems of operator acceptance of technological solutions and difficulties in their ultimate integration for viable operator use.

For example, it has been suggested that when tools dominate, rather than constrain, the joint system, the designer runs a strong risk of solving the wrong problem, and of creating new problems and undermining critical, existing work strategies in the process[ii]. Certainly, the literature provides a number of examples in other domains, including the airline cockpit and process control domains, where failures have been associated with a technology-centred approach to automation at the expense of operator-in-the-loop issues[iii]. This argues for addressing both tool building and tool use if a successful joint operator-machine system is to be achieved. Moreover, failure to do this at the outset, at the conceptual analysis and design stage, or proceeding from a technologically-centred perspective, runs the risk of designing a system that forces operators to adopt procedures and strategies that might in the end degrade, instead of enhance, total performance because of the resulting cognitive dissonance between the operator and the automated system. These remarks are not a criticism of technology per se, but of the failure to appreciate the difficulty of designing truly supportive technology, particularly at the level of aiding the human's cognitive processing.

Another important aspect of this difficulty has to do with the shortage of models of human decision making behaviour, competence and performance that can guide requirements-driven design of decision aids. For example, Endsley observes that despite a concerted thrust to provide military pilots with decision aids through programs like Pilot's Associate, information on how tactical aircraft pilots actually process their environment and make decisions has largely remained anecdotal[iv]. Judging from the literature, the situation with respect to the naval environment is not very different.

It is evident then that a deep understanding of cognitive issues and the nature of the role of the operator in the C2 process, accounting for human capabilities and performance limitations, must provide an important foundation to principled design of decision aids. However, designers are confronted with the problem that despite many recent advances in naturalistic domains[v] directed at understanding human decision-making processes in complex dynamic environments, knowledge in these areas to support system design is still somewhat fragmentary and incomplete. This forces the adoption of a pragmatic approach to the development of practical, viable decision aids, based on a blend of solidly grounded design principles and an informed appreciation of areas where knowledge is limited. This is all the more important in the current DREV project because what is involved is more than a conceptual exploration for the purposes of technology investigation. Rather this ongoing work is aimed at contributing to a specification of a DSS for the mid-life upgrade of the Canadian Patrol Frigate (CPF). A critical constraint is the timing of this upgrade which is expected to take place in phases commencing early in the next century.

This paper examines a range of issues currently being investigated for the design of the DSS, focusing on automation, cognitive and methodological issues. Automation issues address principles and paradigms for automated decision support. Cognitive issues deal with the specifics of the various cognitive-level behaviours which the DSS must exhibit and/or support. The emphasis is on a cognitively-based model of

the Command and Control (C2) process that appears promising for guiding system design. Methodological issues are concerned with managing the complexity of the system design problem in a systematic and effective manner. Some key ideas of a specific model-based framework for design are exposed.

This paper is organised as follows. Section 0 describes the decision-making environment of shipboard command and control. Section 0 describes a wide range of issues related to providing automated decision support in this environment. Section 0 presents key ideas of a model-based framework for support system design. Section 0 briefly describes a model of data fusion and some of its consequences for decision making. Section 0 looks at the domain of naturalistic decision making to identify characteristics of the human decision process in naturalistic environments. A cognitively-based model for the Command and Control process is presented in Section 0. Section 0 discusses the current version of a high-level framework for a decision support system for shipboard command and control. Section 0 provides conclusions.

DECISION-MAKING ENVIRONMENT OF SHIPBOARD COMMAND AND CONTROL

Overview

In a modern frigate like the CPF, most tactical decision making is performed within the ship's Operations Room. There, a team of combat system operators interact with a CCS through consoles, aided by a number of other systems. They perceive and interpret information available from ownship sensors or data-linked from other co-operating platforms, and plan and conduct mission operations. Major C2 tasks include: weapon and sensor systems control, threat evaluation, weapon selection, navigation and ship manoeuvres, and mission planning and evaluation. The C2 process necessitates highly dynamic information flows and decision making involving a number of operators, with a concomitant requirement for developing a common, shared representation of the situation. We now describe various elements of this process, focusing, for brevity, on aspects of situation representation. The principal aim is to expose some aspects of the C2 process model presented in Section 0.

A ship's command structure is typically organised hierarchically. In such a structure, the team of operators is divided into sub-teams, generally along warfare areas, with immediate control exercised by a sub-team supervisor. The Commanding Officer is responsible in all respects. However, effective tactical decision making depends on a co-ordinated team effort and communication among its members is critical in sharing information. Various means are provided to enable this communication and help establish a common situation understanding needed for co-ordinating actions. For example, operators monitor information disseminated to and from other units at sea and ashore, communicate with each other and provide feedback by means of headphones. In addition, stateboards disseminate current information on perceived threats and assist in activating pre-planned responses to highly time-stressed events such as the sudden detection of an anti-ship missile (ASM) flying toward the ship.

Operator behaviour in performing tactical C2 involves a number of demanding perceptual and cognitive processing activities. For example, with respect to the event sequence from "birth" to "death" of a single contact (track), these activities span the moments from first contact detection, its investigation and evaluation in the context of the current mission, development of one or more courses of action, to a course of action decision, potentially involving an engagement and monitoring of that engagement. Information on an unknown contact is obtained from sensors and other sources, possibly involving explicit actions to reduce uncertainties (e.g., send a helicopter for closer surveillance; manoeuvre the ship and observe the contact's response), and this information analysed to evaluate whether it is a neutral, a friend, or a threat. A given contact may undergo several changes in evaluation status over its lifespan. Even if it has been evaluated as a threat, a contact may never be involved in engagement processing by ownship. For example, it may be perceived to be a threat because its behaviour suggests that it is employing stand-off deployment tactics whereby it remains outside the engagement range of ownship's weapon systems. Yet its status and behaviour must be monitored, the implications of its likely intentions understood, and various preparatory decisions and actions taken in readiness for an attack.

Beyond the above snapshot of activities related to a single contact, at any moment numerous contacts may have been detected, with each contact at its own point in its processing sequence. In addition to individual contact processing sequences, operators may need to monitor intercontact relations (functional and spatial relations, priorities, etc.). This is to provide additional clues to a contact's intentions and tactical impact on the mission, permit ranking threats in case of weapon contention, or allow selecting appropriate response actions. As examples, an air platform may have been detected communicating with a surface

contact, perhaps suggesting that its function is to provide targeting information to the surface platform; or a specific engagement action is unwise because the threat is in the path of a friendly contact with a risk of fratricide.

While the above has focused exclusively on problems in the tactical situation that stem from the presence of contacts, there is in fact a variety of other situation elements that may serve to alert the operator to other types of pending problems and which therefore also need to be "tracked" in case a problem does develop. Context independent examples include various numeric or symbolic state variables that reflect the status of the ship's fighting resources (number of missiles and shells remaining, operational and engagement status of sensors and weapons, and so on). Another type of example arises from the need to monitor the effects of a plan for problems in its implementation due to plan execution error, action outcome uncertainty, or unanticipated variability in the external environment. Such problems signal the need for corrective actions. Context dependent examples of problems are specific to situation context. For example, geographical and environmental constraints of the littoral environment can significantly reduce the size of the battlespace or degrade weapon and sensor performance[vi]. This may force operators to react to a new set of problems to the mission that are not of importance in an open-ocean scenario. This can include problems with perception and action (e.g., an increase in number of false alarms by sensors unless the detection threshold is lowered; limits in ship manoeuvrability that impact feasibility of a particular countermeasure).

The common feature of all the above examples is that they relate to the need by operators to understand the meaning and significance of potential problems posed by the ship's dynamic external environment. For our purposes, we define a *problem* to be any event that could negatively impact the achievement of one or more *goals* or suggest, at least, a possible need for change in the way these goals are being, can be, or should be achieved. A problem therefore represents an important goal-relevant property of the environment in that it has the potential for shaping some aspect of an operator's behaviour. We shall return to this observation in Section 0.

Another important type of goal-relevant property for an operator interacting with a complex dynamic environment is related to opportunities. *Opportunities* are defined as events that fortuitously, and usually unexpectedly, present themselves, representing possibilities for achieving goals, or shortcutting their achievement, or for resolving obstacles to their achievement. For example, the appearance of a particular geographical or environmental feature may be an event that offers opportunities for concealing detection from the enemy. In some cases, there may be a cost attached to taking advantage of an opportunity (e.g., manoeuvring the ship from a pre-planned course to take advantage of terrain masking to hide from a threat, which intelligence sources have suddenly signalled, uses fuel but increases the chances of ship survivability). This cost may need to be estimated as a precursor to a decision. We suggest in Section 0 that the three dynamic elements, consisting of *goals, problems* and *opportunities*, along with their dynamic *relations*, can be used in a cognitively plausible triad, to permit structuring a situation representation for tactical decision making that has psychological relevance for the operator.

The demands for perceptual and cognitive processing described above are further increased by the fact that the underlying information is derived by continuously fusing data from a variety of sources, including radar, electronic support measures, infrared search and track, identification friend or foe transponder responses, as well as intelligence information from shore and various deployed units. This information is used to build a coherent Maritime Tactical Picture (MTP) of the ship's area or volume of interest, leading to processing large amounts of data under stringent time constraints. Moreover, the generally imperfect nature of the data, which can be uncertain, incomplete, imprecise, inconsistent, or ambiguous, or some combination of these, due to limited sensor coverage, report ambiguities, report conflicts, or inaccuracies in sensed data[vii], means that at any given moment the MTP only approximates the true state of affairs and that there may be several likely interpretations of the tactical situation. At present, in the CPF for example, these various fusion tasks are performed manually by operators.

Domain Complexity

The AWW environment calls for operators to work effectively in a highly co-ordinated manner toward common objectives. They must: i) continuously scan consoles and monitor communication nets for significant events and alerts; ii) exchange information among themselves or pass information up the chain of command; iii) issue or respond to orders depending on an operator's position and role in the chain of command; and iv) focus attention at any given moment among several competing stimuli and divide

attention between several competing or complementary multiple tasks. In fact, so much needs to be done at any given time that careful attention and time management at both the individual operator and team levels are important for effective performance. Critical incidents (problems or opportunities) can happen at indeterminate times, resulting in dynamically shifting, multiple goals and numerous perceptual and cognitive tasks to be performed by various operators toward co-operatively accomplishing these goals. The result is a complex, dynamic, real-time, data- and goal-driven multi-tasking environment in which goals are continuously created, prioritised and steps taken toward their achievement with continuous attentional shifts between goals. It is therefore vital that both human and machine resources, including weapons, sensors, computers and communications, be effectively managed. Such problems are particularly hard when careful scheduling of shared resources is required to achieve time-constrained goals.

We have described the complexity of this environment in terms of the perceptual and cognitive demands imposed on operators. Woods[viii] states that there are four dimensions that define these demands in a work domain: dynamism, the number of its parts and the extensiveness of interconnections between those parts, uncertainty, and risk. The AWW environment rates as a highly demanding domain in all these dimensions. In practice, complexity can vary, depending on the nature of the conflict. Future combat scenarios are expected to span low to high intensity levels of conflict in open-ocean and littoral areas. In high clutter, terrain-masked littoral environments where there may be many kinds of platforms of many nationalities, with potential for interactions with neutrals becoming embedded within engagements, establishing, identifying the numerous contacts, determining their intentions, interpreting dynamic, complex rules of engagement and taking appropriate action, will be very challenging operator tasks. Complexity also increases with advances in missile technology (e.g., higher speeds, sea skimming attack profiles, smaller signatures) that lead to reduced detection ranges and reaction times against such threats, increases in a ship's region of interest, and leaner operator manning of ships with less people to share the processing load.

Providing Automated Support for Tactical Command and Control

This section examines a number of issues germane to the design of a DSS to support operators in their various roles in the tactical C2 process.

Universe of Solution Approaches

Section 0 highlighted the perceptual and cognitive processing demands of tactical C2 for shipboard combat system operators. These demands vary with the complexity of the situation. While we are concerned in this paper principally with providing automated decision support to help operators meet these demands, it is important to note that there are in fact a number of other approaches, some of which are indicated in Figure 1, for addressing potential decision-related problems in the tactical C2 process. These various approaches can be considered separately or in combination with decision aiding. In fact, the evident interdependencies of these approaches suggests a need for their joint analysis at some stage of the solution process. This analysis would also compare the benefits and costs of the various approaches and establish trade-offs and shortfalls.

- Change Combat System Operator Organization
- Change Doctrine
- Change Tasks
- Change H/M Function Allocation
- Change Manning
- Improve Automation
- Better Selection
- Improve Training

Figure 1 - Universe of Solution Approaches

Automation Issues

Overview

Natural questions to ask concerning the provision of automated aids for improving operational effectiveness of decision making in the Operations Room are: which operator roles and positions in the

Operations Room need to be aided, why, when, and how? Answers to such questions need to be derived based on an appropriate system development philosophy and a coherent design methodology[ix,x].

It is evident from Section 0 that automated support is potentially a highly beneficial option, if not a must, for improving performance in most, if not all, operator positions. For example, in highly dynamic scenarios handling the large amounts of data could quickly overwhelm human capabilities. This also arises from increases in the ship's region of tactical interest that require tracking and understanding the significance of a large (and growing) number of contacts. The fact is that human information processing is subject to a number of limitations and deficiencies, such as finite cycle time, limited working memory, limited ability to perceive and process information and cognitive biases[xi]. It is also negatively impacted by environmental factors or stressors and almost random mistakes (errors of judgement) or slips (errors of execution)[xii].

The form and variety of support would need to be carefully tailored to an operator's position, depending on the nature and mix of the perceptual and cognitive processing involved, and ideally be capable of personal adaptation to suit the variety of support requirements of an operator in that position. The various processes of an operator's role in the decision process would need to be established, decomposed into sub-processes, and decisions made about which of these various sub-processes are candidates for receiving some kind of automated support. An important consideration in making such decisions is the relative capabilities of humans and machines for performing various tasks[xiii,p.84] (e.g., the human is generally considered better at tasks that involve inductive or common-sense reasoning, whereas the machine is better at deductive reasoning). Despite these considerations, we continue in this paper to speak about a single DSS to support operators without differentiating individual operator support requirements.

Aiding Metaphors

An aiding metaphor is concerned with how support is to be provided by a decision aid to the decision maker on some perceptual or cognitive process which is a sub-process of the decision process. The decision aid acting as a prosthesis or as a tool leads to two very different metaphors. In the prosthetic approach, the role of the aid is to replace the operator in some way or to compensate for some human deficiency. The operator is essentially out of the loop and plays a mostly passive role. A frequent criticism of this approach is that it leads to brittle systems, because of limitations in their encoded domain knowledge and assumptions that narrowly bound their view of real-world complexity. This makes them prone to poor performance in the face of environmental variability that has not been anticipated by the system designer. There is an extensive literature in the cognitive engineering[2,3,8] and naturalistic decision making communities[5] which argues instead for the latter approach. In the decision aid as tool metaphor, the aid may be a tool in the hands of a competent but resource limited agent[10]. There is sufficient flexibility, however, for the aid to adapt to a novice, with limited experience (or maybe just a battle-fatigued expert!). Importantly, the operator plays an active role and the tool acts like an intelligent assistant and collaborator. Design emphasis is on supporting the strengths and complementing the weaknesses of the operator. In addition, support is provided for the operator's naturally preferred strategies (at the expense of enforcing a normative or prescriptive approach).

There appears to be a place and need for both metaphors, or some adaptable hybrid of these extremes, in aiding operators, depending on situation context, the specific nature of the processing involved, and the role of the operator. For example, a prosthetic mode would seem appropriate in situations where the operator's current cognitive resources are overwhelmed and he/she is incapable of active, effective participation. There needs, however, to be some understanding by both designer and operator of how the aid's performance itself degrades in such circumstances to avoid the problem of the "blind leading the blind". Also, the minimal involvement of the operator must be determined to avoid, or at least limit, the effects of the "out-of-the-loop performance problem"[xiv] which leaves the operator handicapped in his/her ability to resume control in case of automation failure or once the cognitive demands of the situation have diminished to an acceptable human level. In less demanding situations, the decision aid as a tool mode would keep the operator in the loop.

Need for Design Principles and Guidelines

A number of specific questions are relevant for designing aids for the decision-making process. We give some examples. When environmental tempo and situation complexity increase, leading to cognitive processing overload, what aspects of the environment does the operator continue to need to maintain an

understanding of? If the operator is withdrawn from the loop in stages, what support for maintaining situation awareness should the system provide at each stage to allow the operator to make the judgements and decisions that remain part of his/her responsibility (i.e., not part of the system's)? How should the operator be able to influence the behaviour of support components of a system and how much does the operator want or need to understand about their processing (e.g., models and algorithms used; assumptions made)? How is the authority for deciding the outcome of a supported cognitive process to be distributed between the operator and machine, and how does this depend on the type of process (whose consequences or effects may vary from benign to lethal)?

The need for design principles and guidelines that address these and other questions is evident. Jones and Mitchell[xv] have proposed only general principles on human authority, mutual intelligibility, openness and honesty, management of trouble in case of problems in human-machine communication, and support for multiple perspectives. More specific additional guidance is needed for design.

Delegation of Authority

Section 0 dealt with the issue of how automated support should be provided (i.e., the various aiding metaphors, depending on situation context) once the design decision has been made to provide a decision aid for a given sub-process of the decision process. Some separate, but related, issues concern the mechanism for delegating authority to the system for making the decision on the outcome or result of a supported sub-process, an override capability for the operator when the operator and system have overlapping responsibilities for a sub-process, and the capability for the operator to influence system behaviour when he/she has delegated or lost authority.

Two design approaches to task delegation that appear in the literature are adaptive automation and providing a fixed variety of operator-system modes. Adaptive automation involves an adaptive allocation, depending, for example, on which party has at the moment more resources or is the more appropriate for performing the task[xvi]. A potential problem, however, is that it requires operators to keep up with who is doing what as the allocation changes.

One possible version of the approach of providing various modes of operator-system delegation, which bears some resemblance to that currently implemented in the CPF for threat evaluation and weapon assignment related tasks, is illustrated in Figure 2. Representations are given of five operator-system modes of operation, along with variation in the levels of work distribution and synergy between the automated system and the operator implied by these various modes. Mode selection is made at the system level by the decision maker. The system operates in the selected mode until mode transition is triggered by the decision maker making a new selection. Each mode implies a fixed delegation of authority for all the various sub-processes of the decision process for which automated support is available and use of a fixed support paradigm. The bi-directional arrows for support in Figure 2 indicate that the support paradigm involved in a specific mode could involve the operator in an active role (as in the decision aid as tool metaphor). The actual support paradigm used in a given mode is fixed, but it does not have to be the same in all modes.

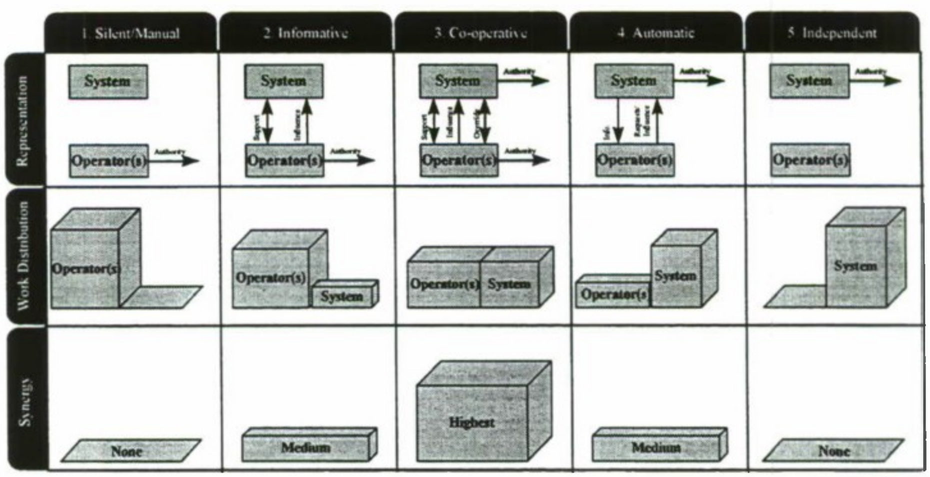


Figure 2 - Operator-System Modes of Operation

The silent/manual mode is characterised by the fact that the operator has total authority and responsibility. Moreover, the system is completely passive and provides no support whatsoever to the operator.

In informative mode, the system only provides the operator with support information (the limit of its work responsibility), some of which may be a consequence of a request from the operator; however, authority again rests totally with the operator. The operator can also influence characteristics of the support provided.

In co-operative mode, the system and the operator work co-operatively. Authority may be divided (e.g., some judgements, decisions and action responsibilities allocated totally to the operator, the rest to the system, depending on type, say) or shared by the two parties. However, in the shared case, one of the two parties (operator or system) has ultimate authority to override the other. This necessitates an override protocol. For example, suppose that the operator decides to retain overriding authority for some types of decisions but is supported in these decisions by the system's processing. Two possibilities for an override protocol in this case are: the system processes, decides and acts accordingly only if the operator first concurs; and the system processes, decides and acts automatically unless the operator vetoes. The operator can also influence the system's processing in sub-processes for which authority has been allocated entirely to the system, for example by requesting that a specific algorithm is used. There is maximum synergy between the system and the operator in co-operative mode.

In automatic mode, the system has total authority, but the operator can influence its behaviour and request information. Apart from such influence, the system can operate in complete autonomy.

The independent mode completely excludes the operator from the process. The system has full authority and responsibility. It processes information and acts autonomously without consulting the operator. In this mode, the system operates as a black box without any required operator interface.

The division of roles between the system and the operator in the various modes is summarised in Table I.

Table I - Operator-System Roles in the Various Modes of Operation

Mode	Operator's Role	System's Role
1. Silent/Manual	Decide and act	Passive
2. Informative	Decide and act Influence system behaviour	Support
3. Co-operative	Decide and act Influence system behaviour Override system	Decide and act Support Override operator
4. Automatic	 Request information Influence system behaviour	Decide and act Provide information Respond to operator influence
5. Independent	Passive	Decide and act

It is also possible to develop an adaptable, hybrid approach encompassing aspects of the two possibilities described above along two dimensions: specific sub-process and situation context. For example, depending only on the specific (type of) judgement or decision, and under specific conditions on the context approved in advance by an operator, the system could make a judgement or decision on his/her behalf. Adaptive, mixed-initiative operator-computer behaviour is also possible in one or both of these dimensions. In the extreme, the system would identify and decide how to satisfy the needs of operators based on some embedded operator model.

Operator Trust

Whatever automation approach is adopted, the operator must have sufficient confidence in the technological solution to use it and be able to delegate his/her authority to judge, decide or act to the

system as and when the need arises. Delegating authority implies that the operator will need a basis for establishing trust in the system[xvii]. General principles for calibrating this trust so that the operator can use an aid discriminatingly and effectively are provided by Muir[17] (i.e., so that the operator does not consistently underestimate or overestimate the aid's capabilities). Design guidance is needed for ways of achieving and maintaining an operator's trust.

TOWARD A METHODOLOGICAL FRAMEWORK FOR DECISION SUPPORT SYSTEM DESIGN

This section describes key ideas of a design framework for a DSS to support operators in their various roles in the tactical C2 process. The methodological approach is suggested by recent work on a theoretical, model-based framework, known as Ecological Interface Design (EID), for designing interfaces for complex human-machine systems[xviii].

Overview

The traditional software engineering approach to system development follows a life-cycle approach involving various activities like requirements gathering and analysis, specification, design, implementation, testing and evaluation, and maintenance[xix]. There are a variety of user participatory refinements (e.g., rapid prototyping[xx]) that relate to how and when the end users of the system are implicated in the design process. A critical problem is that of eliciting users' requirements so that a system is produced that indeed meets these requirements. Wilson and Rosenberg quote statistics to indicate that failures in the requirements phase of software development account for 82% of the cost to fix errors in the final product[20]. Unfortunately, despite the evident good sense of user participation in the process, there is not much evidence in the literature of its successful application in building real support systems. In fact, there is a surprising paucity of examples of operational decision support systems. Moreover, approaches like rapid prototyping are not as rapid as one might expect or want, and many developers feel that existing methods for decision aid development are inadequate[xxi]. Sharp[10] cites a number of problems with user participatory methodologies, including: their essentially empirical trial and error approach; users have a difficult time predicting what they would really like, even if they are expert at what they do; and users' time for involvement in system design is very limited.

Why is the design of a DSS such a difficult problem? We have already touched on some of the issues in Section 0. The reality is that in the absence of an intelligent strategy for investigating a large design space, DSS design is fundamentally an ill-defined problem, likened to solving a jigsaw puzzle consisting of uncertain pieces and an uncertain goal picture, with pieces representing design choices and goal picture the system. Faced with this dilemma, it would appear that the only possible strategy is to engage in many iterative, bottom-up design probes, with continual technology assessment and user evaluation and feedback at each step to direct the search from one prototype to the next. Consistent application of such an approach is problematic. It is potentially very ad hoc, expensive in both time and cost, and can result in much wasted effort. The problem is primarily with the search process which does not incorporate any mechanism beyond user feedback and the developer's intuition to guide the process[10].

Model-Based Methodologies and Ecological Interface Design Theory

Recent work[xxii,xxiii,10] appears to offer a well-founded alternative for developing cognitive support systems. It represents a top-down approach, based on an improved life-cycle that derives power from use of a variety of operator-environment models that effectively help search the design space efficiently.

Not unexpectedly, the power of this model-based approach lies in the choice of models. Some general requirements on operator-environment models to support DSS design are: i) they should have both descriptive and predictive abilities[10]; and ii) they should be operator-centred, i.e., based on knowledge of the operator's processing requirements and their psychological relevance to the operator. Their descriptive abilities allow for understanding current operator-environment behaviour. Their predictive abilities allow the designer to anticipate the consequences of design choices. Item ii is related to our remarks in Section 0 about the need to consider tool use in the analysis stage of producing a DSS. These various models also allow to identify ways in which the system designer can provide support that reduces the processing demands on operators by matching their perceptual and cognitive resources to the demands of the work environment.

Some model-based approaches concentrate on modelling data, information and knowledge needs of the operator[22,23]. However, as Sharp[10] points out, such models alone address only *content* issues: *what* is the information that the DSS could usefully provide to the operator? Additional models are needed to identify:

the *structure* of this information, that is, how the information should be organised to capture relations that are truly significant to the operator (as opposed to the designer); and its *form*, that is, how the information should be mapped onto display features of the interface. The need for these types of models can be traced to two key concepts in EID theory. Although originally introduced as a theoretical framework for designing interfaces for complex human-machine systems, they also help to structure the DSS design problem. Motivated by Vicente and Rasmussen's work[18], these structural prescriptions applied to DSS design can be summarised, along with their model requirements, by:

a) describe the complexity of the external environment in a psychologically relevant way for the operator based on suitable representational models of the environment; and
b) communicate this complexity in an effective manner based on a model of mechanisms people use to reduce the processing demands of environmental complexity.

The connection of EID theory to the three types of issues described previously is that the representational models provided in response to (a) help identify the *content* and *structure* of the information that the DSS needs to provide to the operator, while the model of mechanisms that people use to deal with environmental complexity, referred to in (b), give its *form*.

EID theory has strong motivating connections to Gibson's theory of perception[xxiv] and work in ecological psychology[xxv], with important specific principles as consequences for guiding the design of the interface of the DSS[18]. EID theory currently promotes a specific environmental representation formalism in answer to (a), viz., Rasmussen's Abstraction Hierarchy (AH), while Rasmussen's cognitive control model, known as the skills, rules, knowledge (SRK) framework, provides the answer to (b)[18]. AH is a multilevel representation that describes the various layers of behaviour inducing constraints in the environment. Its power is that it structures the knowledge representation of the environment in a computationally efficient and psychologically valid representation for problem solving to allow the operator to efficiently and quickly cope with unanticipated events, even when they have not been anticipated and designed to be directly supported by the system designer[18]. The SRK framework defines three qualitatively different cognitive levels on which people process information and which the DSS should therefore support to some degree. These levels are based on the operator's degree of familiarity and expertise in dealing with the environment and on the nature of this information which can either correspond to a familiar event, an unfamiliar but anticipated event, or one that is both unfamiliar and unanticipated.

It is worth noting that despite the novelty of EID theory and the fact that it is a very recent development in the field, its technology transfer to industry has already been occurring, primarily in the nuclear and process control industries[xxvi]. It has also been examined recently for its applicability in aviation[26] and the neonatal intensive care domain[10]. In either case, preliminary results established potential for meaningful and useful application. The latter application also drew attention to some of the limitations of the AH as a specific environmental representation for capturing the full set of diagnostic behaviours of physicians. However, this only underlies a need for careful selection of a model structure for the environment. In fact, as our cognitively-based model of the C2 process in Section 0 indicates, there are a variety of cognitive processes of operators that an environmental representation may need to support in applying the framework of this section to the DSS design problem.

Another important aspect of model-based methodologies is that they use field studies to do data collection in situ[10] (i.e., in a realistic work setting like a team trainer). The idea is to observe and record in an exploratory, minimally non-intrusive manner the actual behavioural streams of the work environment for purposes of instantiating the various models identified by (a) and (b). Various structured techniques derived from ethnographic research have been used. The techniques structure verbal protocols according to a conceptual framework or assumptions about the nature of the cognitive activities to guide the data collection. The reader can consult Sharp's work[10] for a review of field study techniques in this vein.

Finally, we note that there are a number of other important issues that need to be addressed at the data collection and work analysis level of a model-based approach to design. For example, we need to be able to identify the various cognitive strategies (how they do what they do), competencies (what they know) and knowledge structures (how they represent what they know) operators employ in processing information. We also need knowledge on how operators deal with work demands and models of how their performance degrades under increasing cognitive demands, and so on. This is to enable identifying ways of providing automated support along the lines of the considerations discussed in Section 0. A framework for these

types of cognitive analysis is suggested by Rasmussen's Cognitive Work Analysis (CWA) framework[xxvii]. CWA is distinguished by its focus on the work domain instead of the more usual focus on tasks in a cognitive task analysis (CTA)[xxviii]. The advantage of CWA over a CTA is that it allows to analyse knowledge-based behaviours of operators in handling unanticipated events for which a pre-planned response is likely to fail, as well as more usual procedural behaviours associated with enacting pre-planned responses.

A Functional Model of Data Fusion

One question that we have not yet broached concerns the specific automated data and information processing technologies which are receiving significant attention in military applications and which are expected to play an important role in next generation military systems for aiding decision makers. A key emerging technology with consequences for decision making is data fusion. Reasons for interest in this technology include the rapid increases in the available data that can be used to compile a tactical picture, leading to huge increases in computational requirements for its production, and the potential for improvements in the tactical picture derived from extended spatial and temporal coverage, increased confidence, reduced ambiguity and improved target detection[xxix].

Based on the work of the Joint Directors of Laboratories (JDL) Data Fusion Subpanel, Waltz and Llinas have defined data fusion as "a multilevel, multifaceted process dealing with the detection, association, correlation, estimation and combination of data and information from multiple sources to achieve refined state and identity estimation, and complete and timely assessments of situation and threat"[xxx]. The process is also characterised by continuous refinements of its estimates and assessments, and by evaluation of the need for additional data and information sources, or modification of the process itself, to achieve improved results. Data fusion is therefore a many layered processing strategy.

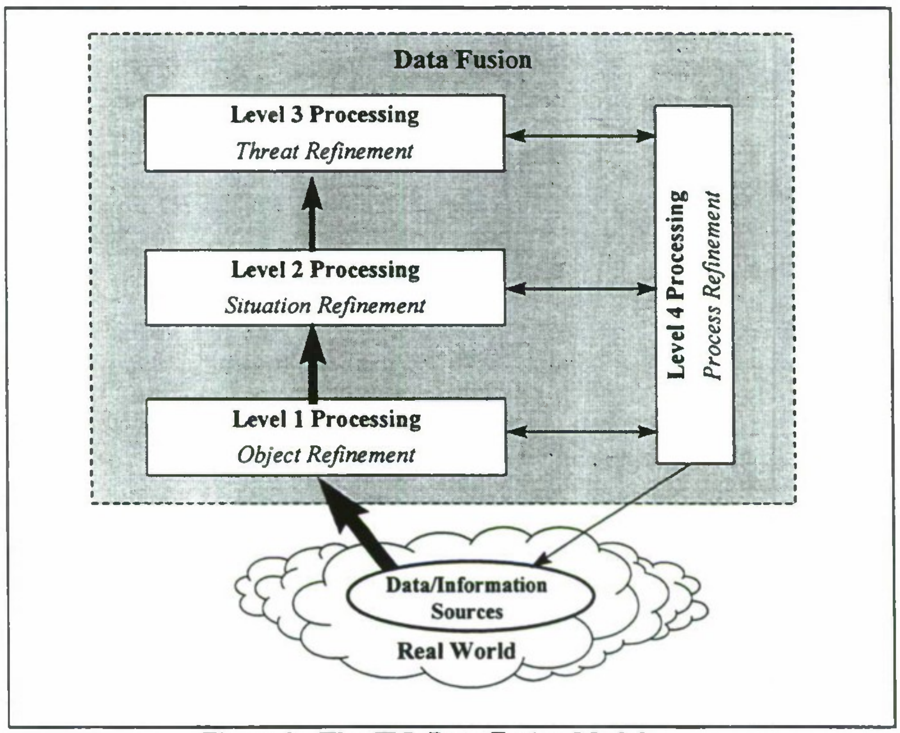


Figure 3 - The JDL Data Fusion Model

Figure 3 shows the four processing levels defined in the JDL data fusion model. In this representation, only multi-source data samples of the real world are available. The interpretation of the real world will be facilitated by data fusion processing. The arrow widths in the figure represent the relative data bandwidths between the processing levels. Each successive level represents a higher level of abstraction and refinement. Level 1 processing corresponds to Multi-Source Data Fusion (MSDF), while Level 2 and Level 3 processing form the basis for Situation and Threat Assessment (STA). In Level 4 processing, inferences

drawn from the data fusion system may be used to select and/or control the input sources, or alter the fusion techniques themselves.

An unfortunate aspect of the JDL data fusion model is that the human role in the process is not evident or even defined. Furthermore, the model provides no formal way of explicitly tying data and information processing capabilities that could be provided by automated data fusion to the perceptual and cognitive demands and decision making requirements of decision makers who are supposed to be the beneficiaries of improvements in the tactical picture. This makes it potentially quite difficult to use the JDL model as a basis for developing an operator support system for shipboard C2, incorporating data fusion processing, that embodies the variety of automated aiding paradigms discussed in Section 0. The cognitively-based model of the tactical C2 process presented in Section 0 is a step toward resolving these problems.

It is worth noting that we use the term resource management (RM) later in this paper to imply the management of both system resources, which are used to provide input or support for processing functionality, and tactical resources, which are used to affect the environment to achieve some tactical or strategic goal. System resources include "base systems" (e.g., CPU, memory, bandwidth) and software processes (e.g., algorithm choices). Tactical resources include weapons (e.g., missiles, guns, tracking and illuminating radars) and navigational mechanisms (e.g., control of vessel speed and direction). In this general sense, therefore, RM extends the Level 4 processing implied by a strict adherence to the JDL data fusion model.

Naturalistic Decision-making Models

We have already described in Section 0 the need in a model-based approach to DSS design for a variety of models of the operator and his/her environment. To this end, an important consideration is the development of models of operator activities in terms of their various perceptual and cognitive processes. This is in fact useful for the entire C2 process. Our focus here is on descriptive models of decision making in the literature that can guide the development of a cognitively-based C2 process model in Section 0.

The characteristics of the decision-making environment of shipboard C2 described in Section 0 match closely the approach taken in the naturalistic decision-making community which is concerned with how human decision makers actually make decisions in complex, real-world settings. Such settings involve ill-structured problems, uncertain, dynamic environments, conflicting, shifting, or ill-defined goals, many action-feedback loops, time constraints, high stakes and pressures, multiple decision makers, and organisational goals and norms. The naturalistic approach emphasises the point "that phenomena observed in complex natural environments may differ substantially from those observed in the laboratory based on decontextualised tasks performed by novices with little stake in the outcomes"[5]. In fact, much of the more traditional, analytically-based decision-making research that appears in the literature has been criticised on this very point, viz., these efforts study human subjects operating in artificially created laboratory settings using normative models to prescribe rational decision-making behaviour on reasonably static tasks. This certainly raises the possibility of the limited representativeness and generalizability of the results of the latter research to the AWW environment.

Some general characteristics of decision making which manifest themselves in some form in all the various naturalistic models[5] can be summarised as follows.

Human decision making is a cognitive process that is triggered in any specific situation by an initial perception of an occurrence in the environment (a cue) that signals a need or opportunity for a decision. Once triggered, decision making involves two cognitive components: situation assessment and selection of a course of action or a response. Once the commitment to a response is made, it is implemented, usually accompanied by monitoring and feedback from the environment.

Situation assessment, the first cognitive component, is an uncertainty reduction process involving judgements needed to extract pertinent information from the uncertain environment. The nature of the situation is interpreted based on the various perceived environmental cues. A number of components of this process are possible, including, continuous attention to and monitoring of cues, diagnosing and interpreting the significance of cues in light of current goals, assessing whether information is adequate for making an interpretation and seeking further information, as may be needed in uncertain situations where there are insufficient, ambiguous, vague, conflicting or contextually uninterpretable cues, and assessing the level of risk and time pressure present in the situation.

Selection of a course of action or a response, the second cognitive component of decision making, extracts a course of action from the judgements made in situation assessment. This involves recognising the

response requirements posed by the situation, identifying options, evaluating their merits in the context of the assessed situation, taking account of the constraints imposed by the situation, and deciding on a response. Klein's model[5], in particular, emphasises the point that expert decision makers in naturalistic settings match the immediate problem situation to a condition in memory and retrieve a stored solution which is then repeatedly evaluated for adequacy in a serial evaluation strategy. This strategy is based on mentally simulating the effects of an option, until one is found that is deemed adequate[xxxi]. The recognitional process of matching solutions to the situation is symptomatic of Rasmussen's rule-based level of cognitive control, indicative of the human's propensity for perceptual processing over more cognitively demanding knowledge-based processing[18].

A Cognitively-based Model of the Command and Control Process

Overview

Command and Control (C2) is defined as the process by which commanders plan, direct, control and monitor any operation for which they are responsible[xxxii]. In the remainder of our discussion, we use the generic term "decision maker" for the single commander in a given C2 centre with the authority to plan and direct operations in that centre. The environment consists of all entities external to the decision maker (e.g., people, machines, databases, weapon and sensor systems). His/her sphere of control may not extend to all of these entities (e.g., threats, neutrals), and the set of entities under his/her control may only be partial or vary with time and context. For example, contrast the situation of a lone ship acting in a single ship operation with that in one involving a federated architecture of ships conducting co-operative engagement tactics to optimise use of the force's fighting resources[xxxiii].

We note that the cognitively-based model of the C2 process presented in this section is expected to be applicable in a wide range of settings, involving one or several operators interacting with a dynamic environment. For example, we anticipate its application in situations from a single operator in front of a console to a team of operators organised hierarchically, as in the shipboard application. In the team setting, two possibilities are: the model is applied at a macro-level to the team with a single decision maker and the various behaviours in the model distributed among the team players; alternatively, a macro-level, network-based process model could be assembled by connecting separate micro-level operator models according to their functional relationships in the team hierarchy. In the latter case, each operator would become the decision maker for his/her nodal model in the network. In team situations where authority for various (types of) decisions can be dynamically delegated (i.e., a dynamic organisational hierarchy), a dynamic, networked-based process model (i.e., either or both of the inter-model links between nodal models and the mappings between processes and people in the team are dynamic) would be needed to represent such a dynamic structure. These observations generalise in a natural way to higher level organisational structures (e.g., task groups).

Due to space limitations, only key ideas underlying our cognitively-based model are sketched in this paper. As already suggested in Section 0, this type of model can be expected to play a valuable role in applying the model-based approach to DSS design. However, to be useful in this endeavour, it is evident from Section 0 that models cannot be ad hoc. They need to adhere to the requirements of psychological relevance in their environmental representations for the operator. They must be consistent with the need of a DSS to communicate, via its interface, domain complexity in a manner consistent with the natural mechanisms that the operator uses to reduce the processing demands of such complexity. In addition, they should reflect the variety of human behaviours that indeed take place (their descriptive ability) or are likely to emerge with automated decision aiding in conducting C2 (their predictive ability). The latter predictive requirement due to the impact of new aiding technologies is an important, and often overlooked, one; for example, it is not generally represented in naturalistic decision-making models[5].

The model of this section was developed as an important step toward responding to these stringent requirements. While model validity undoubtedly remains an important issue, its incorporation of a range of reasonably well accepted cognitive models from the literature provides a well-founded basis for its claim of cognitive plausibility and relevance to design.

There is an abundance of models of the C2 process in the literature. The reader can consult Foster[xxxiv] for an overview of several competing conceptualisations of this process, including the SHOR[xxxv], OODA[xxxvi], MORS[xxxvii] and M/A-Com[xxxviii] models and the Lawson C2 cycle[xxxix]. However, these various models fail to provide several of the characteristics that we are after. Their principal problem, like that of

the JDL model described in Section 0, is that they are not expressed in a language appropriate for easy identification of operators' perceptual and cognitive processes in conducting C2. They also fail to capture some essential elements of human behaviour with their heavy emphasis on data-driven induced behaviour. As observed in Section 0, operators engage in both data- and goal-driven behaviours.

Finally, we summarise some additional features of the military C2 process that have influenced model development in this paper. They include:

a) In the military domain, the C2 process takes place at various command levels and in various phases at each level. There can be a variety of possibilities for the temporal and spatial extents of interest to its decision makers, which may be prioritised for their significance. Furthermore, each situation will have its own requirements on information quality in the various temporal and spatial regions to support the decision making and action execution activities involved. Although our focus is on the shipboard tactical arena, this alone does not justify the development of a totally separate model of human perceptual and cognitive processes in conducting the C2 process in this specific environment. Naturally, the demands for decision support would be expected to vary with setting, but this is an orthogonal consideration. A truly generic process model should therefore readily accommodate the growing need in C2 to inter-operate between the various levels and within the various phases in each level. For example, the model should be compatible with C2 activities that take place at the various phases of pre-deployment of a mission, in-theatre activities, and with actual real-time tactical activities in both a single-ship or force-level context.

b) Another consideration in the above vein is that since the process of C2 can touch a wide range of settings as indicated above, involving one or several people, a truly versatile, cognitively-based model of the C2 process has to permit mappings between processes and people that are one-to-one, one-to-many, many-to-one, or many-to-many. In view of the multitude of possibilities for concurrency of the various behaviours, a sequential structure on behavioural interactions would be inadequate. In fact, there are good arguments why even in the case of a single operator a sequentially-based process model would lack the required features. Work by Bainbridge provides arguments in this direction[xl].

c) Surprisingly little appears to need being changed in our C2 model at the structural, process decomposition level to accommodate the directions established in (a) and (b). However, the nature of the specific cognitive processing can certainly be impacted. For example, in pre-deployment and in-theatre operations where there is more time for mission-level planning and determination or tailoring of pre-planned responses to be used in the various tactical operations themselves, we would anticipate more evidence of the knowledge-based level of cognitive processing (as defined in the SRK framework), reflective of more anticipatory (and therefore less reactive), pro-active planning behaviours[xli]. We also expect that such processing is already present in the tactical arena itself when, for example, unanticipated variability in the environment (as must be expected in any hostile situation) forces a pre-planned response to be adaptively repaired online before it is implemented. It is certainly the case that in this environment, in view of its complexity and the need to establish common intent among command personnel, heavy emphasis is placed on established doctrine concerning pre-planned tactics in case a ship/force-protected asset is suddenly attacked[xlii]. We would anticipate, however, that knowledge-based processing by operators in the tactical environment will further emerge with the presence of automated decision aids that permit, for example, "optimising" online the use of fighting resources. In view of these various considerations, the C2 model includes both rule-based and knowledge-based behaviours. This is different from the purely recognitional type of behaviour that largely dominates naturalistic models of decision making[5].

d) Current C2 models focus on contacts in the external environment and their status. This is not surprising given the underlying reasons for C2 in the military domain. However, as we observed in Section 0, there are a variety of other potential problems that are likely to be encountered in conducting C2.

e) An important aspect of human behaviour in dynamic environments is that it is opportunistic. Humans have a creative ability for identifying opportunities in a situation and taking advantage of them. This needs to be reflected in a cognitively-based C2 model so that designers can determine if and how it should be aided.

Structuring the Decision-Making Environment

We sketch here a framework for dynamically structuring an environment in a manner that, on the grounds of cognitive plausibility at least, appears psychologically relevant to a decision maker operating within the environment. The need for psychological relevance has already been discussed. The framework represents a personalised structuring of the environment, *from some frame of reference considered relevant by the decision maker*. For example, the frame of reference could be his own, a participating unit's, his enemy's, and so on. It is based on the concept that what the decision maker would probably want to *comprehend* about the environment at any given time to be able to make decisions about action can be phrased in the language of the *problems* or *opportunities* posed by the environment at that time, given the *goals* and the state of various *relations* that the decision maker deems relevant among these problems, opportunities and goals at that time. In short, they are specific, time-dependent, goal-relevant properties of the environment that shape the decision maker's action management behaviour. We refer to these as *situation structures*. We have chosen this term in place of *situation assessments* to avoid the clash that exists between the term situation assessment as used in the naturalistic decision-making literature[5] and the JDL terminology in data fusion[7] that differentiates between situation assessments and threat assessments. In addition, calling them structures emphasises their functional role: they provide a structured representation for understanding the environment as a prerequisite for (rule-based or knowledge-based) action.

The decision maker could be interested in a variety of types of relations among goals, problems and opportunities, including enabling relations (between an opportunity and a goal), causal or subset relations (between pairs of problems or opportunities), value relations (for prioritising goals, problems and opportunities), impediment relations (between a problem and a goal), and so on.

Definitions of the various terms used above have previously been given in Section 0. The motivation behind this general structuring framework has also been illustrated there in the context of the shipboard setting. We note that the separation of events into insignificant events, problems and opportunities, indicated in Figure 4, is not really an event partition. An event could represent both a problem and an opportunity. For example, the presence of a particular geographical or environmental feature in a ship's vicinity, which route planning could avoid, might represent a problem (reduced sensor detection envelope) for achieving one of the decision maker's goals (optimise detection) but an opportunity (increased chance of concealment) for another (optimise survivability). This arises from conflicting goals. There is also a potential for duality between problems and opportunities. For example, a problem in one frame of reference (e.g., his own) could represent an opportunity in another (e.g., his enemy's). This structuring in a variety of frames of reference should be an important element of a decision maker's need for simultaneous, multiple perspectives in understanding the situation in some cases as a precursor to making a decision.

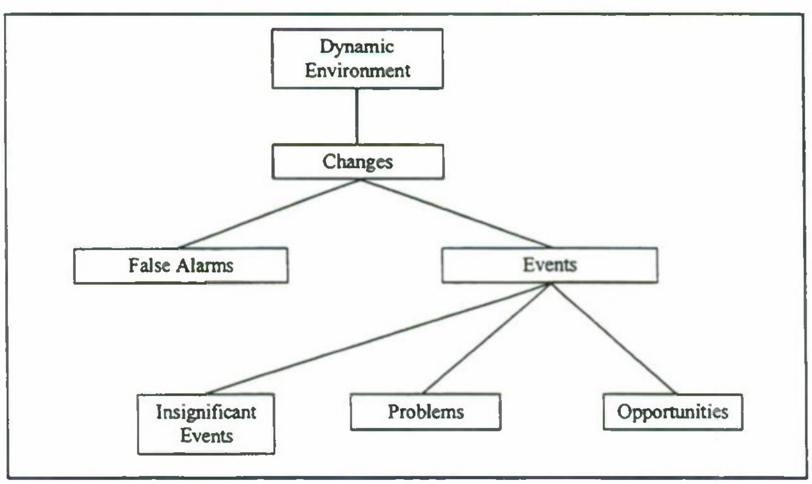


Figure 4 - Identifying Problems and Opportunities by Sensing Change

This representational structuring of situation understanding can be thought of as a dynamic triad of psychological relevance to the decision maker. It is illustrated in Figure 5. Although the triad shown there seems to have a flat relational structure, it is not difficult to see that there are the usual advantages that come from abstraction (e.g., computational efficiency, iterative refinement) for developing the triad, or

portions of it, into an abstraction hierarchy. This raises the question of the form of a psychologically relevant framework for such a hierarchy. Something similar to Rasmussen's AH for means-ends relations[18] naturally comes to mind. However, it is important to note that in the setting here the abstraction hierarchy would be dynamic. We do not pursue this matter further in the present paper.

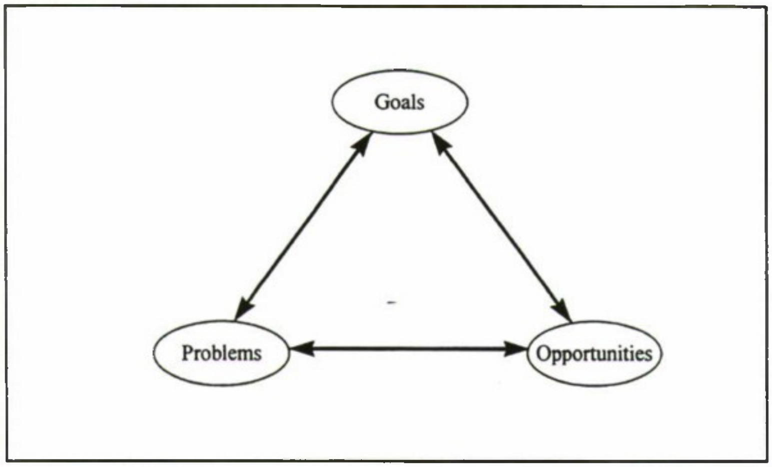


Figure 5 - Dynamic Structuring of Goals, Problems and Opportunities

Finally, we mention that no claim is made here that the decision maker wants or needs to be aware of *all* elements of the dynamic situation structure shown in Figure 5 at any given moment for successful performance in his/her environment. In fact, it is surely possible that the decision maker is not actually aware of *all* these elements at any one time and is still able to achieve quite satisfactory performance. This touches on the larger question of what exactly the link is between situation awareness and performance[xliii].

Description of the Model

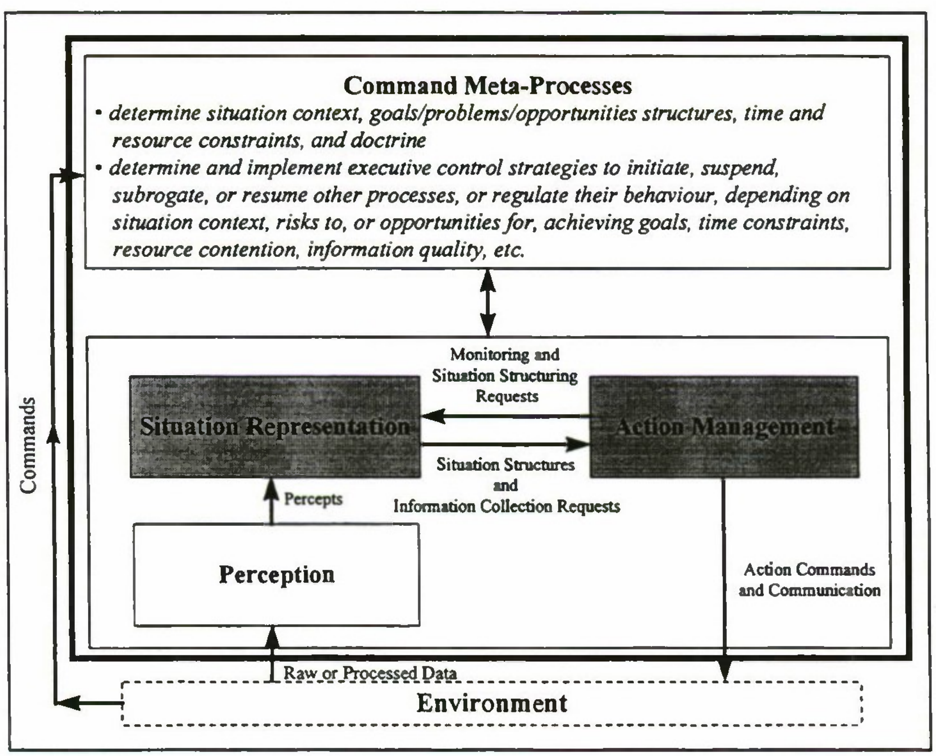


Figure 6 - Command and Control Process Model

We now present our cognitively-based, behavioural model of the C2 process. The model decomposes the C2 process into two levels: a lower level involving the three processes *Perception*, *Situation*

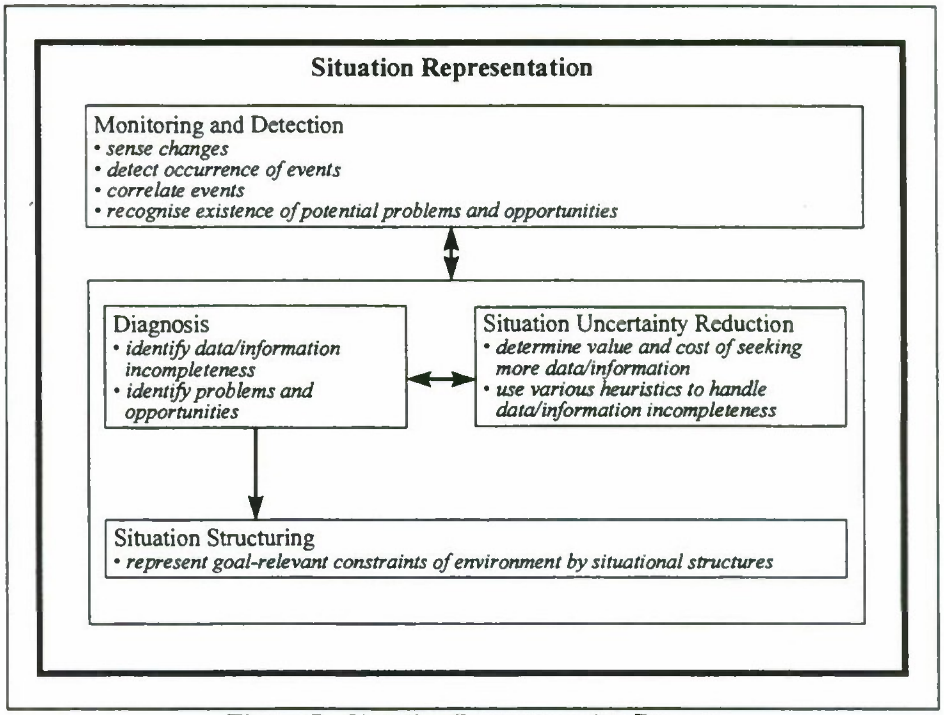


Figure 7 - Situation Representation Process

Representation and *Action Management*, and a higher level consisting of various *Command Meta-Processes*. The details of these various processes are explained in almost self-explanatory manner in three figures, Figure 6, Figure 7, and Figure 8. We now examine some of the highlights of the model.

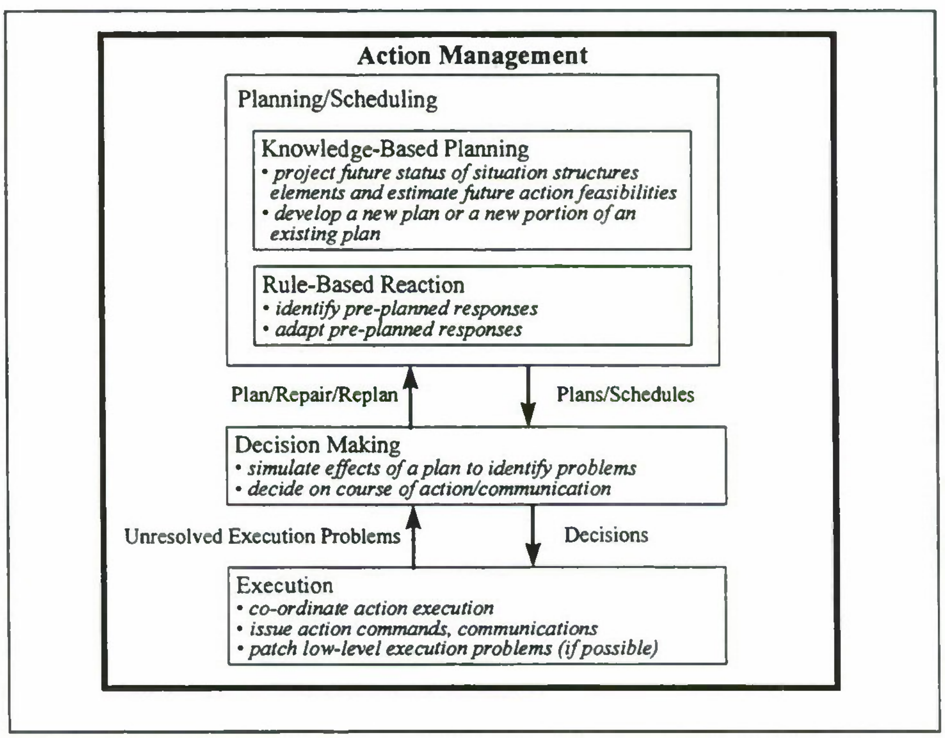


Figure 8 - Action Management Process

The Command Meta-Processes, shown in Figure 6, dynamically manage the goals and choice of situation structures, and control various parameters (like frequency with which to monitor for changes in a specific environmental feature) and the sequencing and cognitive level (in the sense of the SRK taxonomy)

of lower level processes. Feedback from the lower level can cause new goals and situation structures to be generated and old goals and structures to be removed from consideration, as well as new control strategies to be employed. Commands from higher-level command echelons originating in the environment can also induce changes in these processes.

Situation Representation (Figure 7) and Action Management (Figure 8) are uncertainty reduction processes, but in different senses. The first reduces uncertainty in understanding the situation. In this case, it involves judgements about where there is uncertainty, incompleteness, imprecision, inconsistency, or ambiguity, or some combination of these in data/information (short-term knowledge) and in the resolution of such uncertainty. In cases where it is deemed worthwhile to reduce this uncertainty by obtaining more data/information, it can issue an information collection request to Action Management (e.g., send a helicopter for closer surveillance; manoeuvre the ship and observe the contact's response). Action Management reduces uncertainty in the selection of actions.

In either case, it is possible to add the need to reduce uncertainty due to incomplete long-term knowledge (purposeful learning behaviour), by first evaluating benefits and costs of seeking this knowledge and the likelihood of being successful in doing so. Situation Representation would issue a request to Action Management which would handle its own needs and those of Situation Representation with subsequent feedback to Situation Representation. However, these various considerations are not fleshed out in the figures shown.

Situation Representation is the process of producing or generating abstract descriptions or representations of a dynamic environment. The particular descriptions are in terms of situation structures as defined in Section 0 which the decision maker (in his/her Command Meta-Processes) dynamically determines to be relevant for determining and managing action. Situation Representation generates these structures either at the request of Command Meta-Processes or at the request of Action Management when it needs understanding about specific situation elements for planning or for managing the co-ordination and execution of its action decisions. Situation Representation is analogous to the human situation assessment process described in the naturalistic decision-making literature[5]. It combines the situation assessment and threat assessment processes of the technologically-centred JDL data fusion model[30], but from the perspective of the human. The reasons for introducing the new terminology are similar to those previously stated for situation structures. It is also useful to distinguish between process and product. Another important nuance, which distinguishes our approach from previous efforts, has to do with the way the model handles situation projection. While a given situation element (goal, problem, opportunity, relation) may well involve an aspect of the future (e.g., the problem related to a contact might be "Time to ship intercept is less than 30 seconds"), we use the term projection in an action-oriented sense in that its need is determined within Action Management to support knowledge-based planning.

Process sequencing in Situation Representation essentially follows the identification strategy suggested by Figure 4, with feedback loops arising from the need to resolve problems of incomplete data, information or knowledge.

Action Management handles all processes related to determining feasible courses of action, action selection, and management of action execution. Actions commands are commands to physical actuators (sensors, weapons, navigation, etc.) or lower levels in the organisational hierarchy. We also include as part of Action Management decisions related to sharing information with other parties in the environment. This leads to the communication shown in Figure 6 (e.g., a speech turn or set of words spoken to another person in the decision maker's environment; information data-linked to a participating unit or a shore-based C2 centre, etc.). Another reason for communication is to request additional information/knowledge to be supplied from an entity in the environment.

Finally, note the presence of both rule-based and knowledge-based processing in Action Management for reasons previously given in Section 0. This is also the case for Situation Representation. For example, diagnosis could be entirely rule-driven or employ, in addition, various knowledge-based heuristics to make judgements that reduce uncertainty in situation understanding. Evidence of both types of behaviour have been found in the naturalistic decision-making literature. For example, Kaempf, Wolf and Miller[xliv] report findings of a study in which they analysed results of interviews based on use of the critical decision method[xlv] to identify the primary situation diagnosis strategy used by anti-air warfare officers in the Combat Information Centre of an AEGIS cruiser as essentially rule-based feature matching. This strategy consists of matching existing cues with a remembered set of cues. However, in situations of insufficient

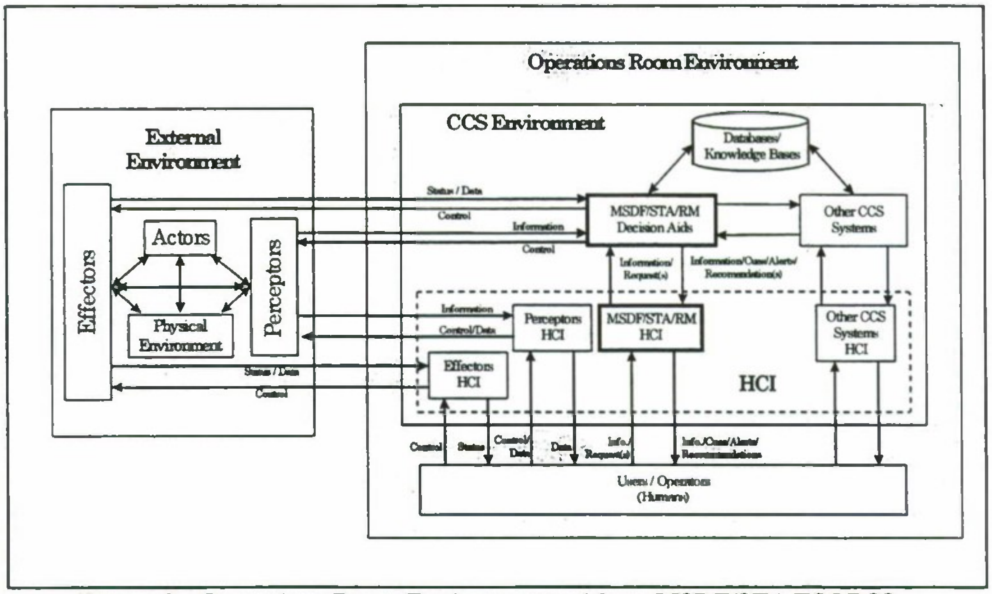


Figure 9 - Operations Room Environment with an MSDF/STA/RM DSS

information or when the situation was novel and unfamiliar, the officers used a knowledge-based strategy of story generation in which the information available is used to build an explanatory story of the situation.

CURRENT HIGH-LEVEL FRAMEWORK OF A REAL-TIME DECISION SUPPORT SYSTEM

At present, in the CPF, the various data fusion tasks that are used to build the tactical picture of the environment external to the ship are manually performed by operators, communicating among themselves. There is some automated support in the CCS for reactive action management related to the allocation of the fighting resources (weapon allocation) in terminal engagement. Automated situation representation capabilities are limited essentially to threat evaluation in the form of threat ranking. As a simple example, the capability for an operator to ask the CCS to monitor a specific contact or group of contacts for a certain potentially threatening behaviour (either pre-defined or defined on the fly by the operator), and alert him/her if such a behaviour occurs, does not exist. In addition, explicit situation representation in the human-computer interface (HCI) is limited to level of contacts. General situation representation as defined in this paper is done in the heads of operators. In theory at least, there is therefore much scope for introducing many of the advanced DSS concepts raised in this paper to the operational environment of the CPF.

DREV's work is expected to lead to a specification of a DSS to support operators in: i) the fusion of data from the ship's sensors and other sources; ii) the formulation, maintenance and display of an accurate dynamic situation picture, leading to enhanced situation awareness; iii) the identification and selection of courses of action in response to anticipated or actual threats to the mission; and iv) action implementation once a decision to act has been made and is being carried out. With respect to particular DSS capabilities, item i relates to its Multi-Source Data Fusion (MSDF) capability. It supports perception activities in Figure 6, for example by enhancing the quality and coverage of the processed data that feeds perception. Item ii relates to its Situation and Threat Assessment (STA) capability. It supports the situation representation process in Figure 7. Finally, items iii and iv relate to its Resource Management (RM) capability. It supports the action management process in Figure 8.

It is envisaged that this DSS, which we refer to as an MSDF/STA/RM system, will become an embedded component of the ship's combat system, integrated within the CCS. A rough, high-level perspective of this integration is shown in Figure 9. The environment in Figure 9 is decomposed into the portion within the ship's Operations Room and the portion outside, including the perceptors (organic and non-organic information sources), effectors (active and passive weapon systems), the actors (threats, friends, neutrals) and the physical environment external to the ship. The CCS environment is everything in the CCS of a hardware or software nature, including the various HCIs, databases/knowledge bases and other CCS systems. These databases/knowledge bases contain a variety of a priori knowledge, including

standard operating procedures and pre-planned tactical responses, and strategic, Electronic Warfare (EW) and intelligence information.

CONCLUSIONS

This paper examined a range of issues currently being investigated for the design of a decision support system to support combat system operators of a modern frigate in their tactical decision making and action activities as part of the Command and Control process. Automation, cognitive and methodological issues were highlighted.

Fundamental issues in providing automated support are related to the questions of which operator roles and positions need to be aided, why, when, and how. These are very complex questions that require an appropriate system development philosophy to be established and a coherent design methodology to be followed if a joint system, comprised of both operators and automated decision aids, is to lead to improved operational effectiveness in conducting shipboard Command and Control. This paper suggests that the recent emergence of model-based frameworks for design offers a significant potential for rescuing the design process from falling into the trap of pursuing an ad hoc approach with high risk for incurring large expense in time, cost and wasted effort. Undoubtedly the dilemma of fragmentary and incomplete understanding of the design process still remains. Faced with these considerations, we now need to turn the various insights offered by what is known into a pragmatic approach to the development of practical, viable decision aids, based on a blend of solidly grounded design principles and an informed appreciation of areas where knowledge is limited. This paper represents only a step toward developing this pragmatic approach.

Ongoing and future work is aimed at developing such an approach. Other work is related to refining the cognitively-based process model for the Command and Control process and examining the implications of this model. While it was derived from thinking about the naval problem, it appears to have wide applicability to a number of other military and non-military settings. This is probably not very surprising since human behaviour and strategies for coping with complexity in a variety of dynamic environments are likely to share much commonality. Like the ant on the beach in Simon's parable[xlvi], detailed aspects of behaviour arise from the impact of the environment. In fact, it is tempting to conjecture that there are equivalence classes of environments in which human behaviour, and therefore its need for support, as well as the nature of effective support, impact support design similarly regardless of the specific member of that class.

In another direction, thinking about the notion of models that have psychological relevance to the operator certainly helped the author better appreciate the immense difficulty (if not impossibility) of a system designer anticipating all variabilities in a complex dynamic environment. Consider, for example, the problem of designing to support the operator in situation representation. Choosing even a representative set of situation structures that would be needed for effective performance in a given situation appears to be a difficult task. The capability for the operator to create and customise his own situation structures on the fly is therefore likely to emerge as an important design consideration.

Finally, work is needed to establish the relation of the present effort to other work in the literature on tactical decision aiding[xlvii,xlviii,xlix] and situation and threat assessment[l,li]. Certainly, an important difference from previous efforts is the emphasis in the current work on developing a principled, holistic approach, encompassing modelling methodologies and cognitive and environmental models, as a formal prerequisite to design. This is in opposition to an approach which does not offer any specific signposts on how to search the design space but rather rests on ad hoc methods and the developer's intuition[10].

ACKNOWLEDGEMENTS

The author wishes to thank his colleagues in the Data Fusion and Resource Management Group, Defence Research Establishment Valcartier, collaborators in the R&D Department at Lockheed Martin Canada, and LCdr Richard Flower, SO TRG C3I, Canadian Forces Maritime Warfare Centre, for many helpful discussions and inputs, as well as comments on an early version of some of the material in this paper.

REFERENCES

1. T. Liang, "Getting the Act Together: Hardkill-Softkill Co-ordination for Littoral Waters", *Jane's International Defence Review*, Vol. 28, pp. 54-56, 1995.
2. D.D. Woods, "Cognitive Technologies", *The AI Magazine*, Vol. 6, No. 4, pp. 86-92, 1986.
3. D.D. Woods. and E.M. Roth, "Cognitive Systems Engineering", in *Handbook of Human-Computer Interaction*, M. Helander, ed., pp. 3-43, Elsevier Science Publishers B.V., North-Holland, 1988.
4. M.R. Endsley and R.P. Smith, "Attention Distribution and Decision Making in Tactical Air Combat", *Human Factors*, Vol. 38, No. 2, pp. 232-249, 1996.
5. G.A. Klein, J. Orasanu, R. Calderwood and C.E. Zsambok (Eds.), *Decision Making in Action: Models and Methods*, Ablex, Norwood, New Jersey, 1993.
6. J.F.G. Wade, "Navy Tactics, Doctrine, and Training Requirements for Littoral Warfare", M.Sc. Dissertation, Naval Postgraduate School, Monterey, California, 1996.
7. E.L. Waltz and D.M. Buede, "Data fusion and decision support for command and control", *IEEE Trans. Syst., Man, Cybern.*, Vol. SMC-16, No. 6, pp. 865-879, 1986.
8. D.D. Woods. and E.M. Roth, "Cognitive Systems Engineering", in *Handbook of Human-Computer Interaction*, M. Helander, ed., pp. 3-43, Elsevier Science Publishers B.V., North-Holland, 1988.
9. H.A. Heathfield and J. Wyatt, "Philosophies for the Design and Development of Clinical Decision Support Systems", *Methods of Information in Medicine*, Vol. 32, pp. 1-8, 1993.
10. T.D. Sharp, "Progress Towards a Development Methodology for Decision Support Systems for Use in Time-Critical, Highly Uncertain and Complex Environments", Ph.D. Dissertation, Department of Electrical and Computer Engineering, University of Cincinatti, 1996.
11. C.D. Wickens and J.M. Flach, "Information Processing", in *Human Factors in Aviation*, E.L. Wiener and D.C. Nagel, eds., Academic Press, Inc., 1988.
12. D.A. Norman, "Categorization of Action Slips", *Psychological Review*, pp. 1-15, 1981.
13. B. Schneiderman, *Designing the User Interface: Strategies for Effective Human-Computer Interaction*, Second Edition, Addison Wesley Publishing Company, 1992.
14. M.R. Endsley and E.O. Kiris, "The Out-of-the-loop Performance Problem and Level of Control in Automation", *Human Factors*, Vol. 37, No. 2, pp. 381-394, 1995.
15. P.M. Jones and C.M. Mitchell, "Human-Machine Cooperative Problem Solving: Theory, Design, and Evaluation of an Intelligent Associate System", *IEEE Transactions on Systems, man, and Cybernetics*, Vol. 25, No. 7, pp. 1039-1053, 1995.
16. W.B. Rouse, "Human-Computer Interaction in the Control of Dynamic Systems", *Computing Surveys*, Vol. 13, No. 1, pp. 71-99, 1981.
17. B.M. Muir, "Trust Between Humans and Machines, and the Design of Decision Aids", in *Cognitive Engineering in Complex Dynamic Worlds*, E. Hollnagel, G. Mancini, and D.D. Woods, eds., Academic Press, Harcourt Brace Jovanovich, Publishers, London, pp. 71-83, 1988.
18. K.J. Vicente and J. Rasmussen, "Ecological Interface Design: Theoretical Foundations", *IEEE Transactions on Systems, Man, and Cybernetics*, Vol. 22, No. 4, pp. 589-606, 1992.
19. P.A. Laplante, *Real-Time Systems Design and Analysis: An Engineer's Handbook*, IEEE Computer Society Press, Los Alamitos, California, 1993.
20. J. Wilson and D. Rosenberg, "Rapid Prototyping for User Interface Design", in *Handbook of Human-Computer Interaction*, M. Helander, ed., pp. 859-875, Elsevier Science Publishers B.V., North-Holland, 1988.
21. NATO, "COADE: A Framework for Cognitive Analysis, Design and Evaluation", Technical Report AC/243 (Panel 8), TR/17, Defence Research Group, NATO HQ, Brussels, 1995.
22. G.A. Sundstrom and A.C. Salvador, "Integrating Field Work in System Design: A Methodology and Two Case Studies", *IEEE Transactions on Systems, Man, and Cybernetics*, Vol. 25, No. 3, pp. 385-399, 1995.
23. D. Fafchamps, C.Y. Young and P.C. Tang, "Modelling Work Practices: Input to the Design of a Physician's Workstation", in P.D. Clayton, ed., *Proceedings of the Fifteenth Annual Symposium on Computer Applications in Medical Care*, Washington, DC: American Medical Informatics Society, pp. 788-792, 1991.

24. J.J. Gibson, *The Ecological Approach to Visual Perception*, Lawrence Erlbaum Associates, Publishers, Hillsdale, NJ, 1986.
25. K.J. Vicente, "A Few Implications of an Ecological Approach to Human Factors", in *Global Perspectives on the Ecology of Human-Machine Systems*, J. Flach, P. Hancock, J. Caird and K. Vicente, eds, pp. 54-67, Lawrence Erlbaum Associates, Publishers, Hillsdale, NJ, 1995.
26. N. Dinadis and K.J. Vicente, "Application of Ecological Interface Design to Aviation", CEL 96-07, Cognitive Engineering Laboratory, University of Toronto, 1996.
27. J. Rasmussen, *Information Processing and Human-Machine Interaction: An Approach to Cognitive Engineering*, North-Holland, New York, 1986.
28. K.J. Vicente, "Task Analysis, Cognitive Task Analysis, Cognitive Work Analysis: What's the Difference?", in *Proceedings of the Human Factors and Ergonomics Society 39th Annual Meeting*, pp. 534-537, 1995.
29. D.L. Hall, "An Introduction to Multisensor Data Fusion", *Proceedings of the IEEE*, Vol. 85, No. 1, pp. 6-23, 1997.
30. E. Waltz and J. Llinas, Multisensor Data Fusion, Artech House, Inc., Norwood, MA, 1990.
31. G. Klein and B.W. Crandall, "The Role of Mental Simulation in Problem Solving and Decision Making", in *Local Applications of the Ecological Approach to Human-Machine Systems*, P.Hancock, J. Flach, J. Caird and K. Vicente, pp. 324-358, Lawrence Erlbaum Associates, Publishers, Hillsdale, NJ, 1995.
32. Joint Doctrine for Canadian Forces Joint and Combined Operations (1995-04-06).
33. Naval Forces Update, "Co-operative Engagement and Self Defence", *Jane's Defence Weekly*, pp. 32-34, 19 June, 1993.
34. G.D. Foster, G.D., "Contemporary C^2 Theory and Research: the Failed Quest for a Philosophy of Command", Defence Analysis, Vol. 4, No. 3, pp. 201-228, 1988.
35. J.G. Wohl, "Force Management Decision Requirements for Air Force Tactical Command and Control", *IEEE Transactions on Systems, Man, and Cybernetics*, Vol. SMC-11, No. 9, 1981.
36. G. Hart and W.S. Lind, America Can Win: the Case for Military Reform, Adler and Adler, Bethesda, Maryland, 1986.
37. Military Operations Research Society, Command and Control Evaluation Workshop, Military Operations Research Society, Alexandria, Virginia, 1985.
38. M/A-Com Linkabit, Inc. National Defense University (NDU) Command and Control Research Program (CCRP), Technical Report 158, M/A-Com Linkabit, Inc., Vienna, Virginia, 1985.
39. J.S. Lawson, "The State Variables of a Command and Control System", in J. Hwang, ed., *Selected Analytical Concepts in Command and Control*, pp. 61-83, Gordon and Breach, New York, 1982.
40. L. Bainbridge, "The Change in Concepts Needed to Account for Human Behaviour in Complex Dynamic Tasks", in *Proceedings of 1993 International Conference on Systems, Man, and Cybernetics*, pp. 126-131, 1993.
41. J.-M. Hoc, *Cognitive Psychology of Planning, London*, London, Academic Press, 1988.
42. S.A. Hyer, J.J. Johnston and C.L. Roe, "Combat System Effectiveness Modeling to Support the Development of Anti-Air Warfare Tactics", *Johns Hopkins APL Technical Digest*, Vol. 16, No. 1, pp. 69-82, 1995.
43. M.R. Endsley, "Toward a Theory of Situation Awareness in Dynamic Systems", *Human Factors*, Vol. 37, No. 1, pp. 32-64, 1995.
44. G.L. Kaempf, S. Wolf and T.E. Miller, "Decision Making in the AEGIS Combat Information Center", in *Proceedings of the Human Factors and Ergonomics Society 37th Annual Meeting*, pp. 1107-1111, 1993.
45. G.A. Klein, R. Calderwood and D. Macgregor, "Critical Decision Method for Eliciting Knowledge", *IEEE Transactions on Systems, Man, and Cybernetics*, Vol. 19, No. 3, pp. 462-472, 1989.
46. H.A. Simon, *The Sciences of the Artificial*, 2nd edition, Cambridge, MA, MIT, 1981.
47. S.G. Hutchins and Kowalski, J.T., "Tactical Decision Making Under Stress: Preliminary Results and Lessons Learned", *Proceedings of the 10th Annual Conference on Command and Control Decision Aids*, Washington, D. C., pp. 85-96, 1993.

48. S.G. Hutchins, S.G. and B.K. Rummel, "A Decision Support System for Tactical Decision Making Under Stress", *Proceedings of the First International Conference on Command and Control Research and Technology*, Washington, D. C., pp. 163-173, 1995.
49. S.G. Hutchins, "Principles for Intelligent Decision Aiding", Naval Command, Control and Ocean Surveillance Center, RDT&E Division, Technical Report 1718, 1996.
50. J.A. Ballas, D.C. McFarlane, L.B. Achille, J.L. Stroup and S.D. Kushnier, "Interfaces for Intelligent Control of Data Fusion Processing", Naval Research Laboratory, NRL/FR/5513-96-9806, April 1996.
51. S. Paradis, B.A. Chalmers and P. Bergeron, "Towards a Joint Human-Machine Model for Situation and Threat Assessment", *this proceedings*, 1997.

This page intentionally blank.

NAVIGATIONAL AWARENESS USING 3-D MAPS; ELEVATION ANGLE, COMPLEXITY AND FEATURE EFFECTS

Joseph C. Hickox, Capt, USAF
Dept of Behavioral Sciences and Leadership
USAF Academy, Colorado

Dr Christopher D. Wickens
Aviation Research Lab, University of Illinois
Savoy, Illinois

Introduction

Fundamental to the skills of aviators is the ability to always know where they are and where they are going. To a novice, viewing the world from the air is a very unique sight. It is often difficult to recognize or even identify known features and landmarks. Air navigation creates a unique set of cognitive demands as a pilot or navigator repeatedly compares features of the map with the outside view or forward field of view (FFOV), and quickly determines whether or not the two are congruent (Aretz, 1991). In this task, referred to as "navigational checking" (Wickens, Schreiber & Renner, 1994), features in the FFOV that are compared with the map could be single entities, or an aggregation of features, such as road intersections, a group of buildings or a series of river bends, etc. The mental processing and cognitive requirements imposed on the navigator engaged in navigational checking are complex and require further examination in an effort to model the variables involved.

The need for accuracy in this task is self-evident, as getting "lost" in the air can have dire consequences (Williams, Hutchinson & Wickens, 1996). Likewise, speed should also be given a high priority in the determination of congruence or incongruence. As an example, consider a fighter aircraft on the final leg of a bombing run over hostile unfamiliar enemy territory. As the aircreaft approaches the objective area, the pilot needs to identify the target rapidly to program the weapons guidance systems. The cost of misidentifying the target can not be overstated as the consequences of bombing the wrong target can be disastrous. During the Persian Gulf War, several highly publicized examples of bombing wrong targets resulted not only in civilian loss of life, but also losses to our own troops in so called "friendly fire" incidents (Bond, 1991; Hackworth, 1991). At the same time, with the speeds that some modern fighters can attain, it is conceivable that an aircraft would only be in the target zone for a few seconds. Identification of the target needs to be timely and accurate. Although many of the guidance systems in today's modern aircraft are automatic, guided by computers, internal navigation, and global positioning, target confirmation and engagement still depends upon the pilot visually identifying the target; and pilot performance will always remain critical should automation fail.

The nature of navigational checking varies with the aircraft's particular phase of flight. The pilot of an aircraft in cruise flight (usually quite high) will be interested in several different viewing elevation angles in the FFOV. The pilot may be looking almost straight down (-90°), or several miles ahead of the aircraft, resulting in lower viewing elevation angles. In contrast, the pilot of an aircraft on a supply airdrop, bombing run or final approach at altitudes below 500 ft., would be less interested in what was at large viewing angles directly beneath, or at steep elevation angles in front of the aircraft and more interested in looking ahead of the aircraft, along smaller viewing angles (i.e., more parallel to the ground).

The Design Problem

It is important to reiterate the type of meteorological conditions that are assumed to exist within the paradigm of the navigational checking task. Although it is unlikely that any pilot would venture into the air without some form of radio navigational aid available to assist in navigation, the paradigm at present assumes that navigational checking is strictly a visual task, accomplished under visual flight rules (VFR). This assumption becomes most relevant during less standardized missions into areas that are not well known or well serviced by ground navigational aids, such as military combat sorties, medical evacuations, and search and rescue missions. In these types of scenarios, time is critical and consequences of navigational errors are severe.

With the advent of modern computerized navigational systems and precision guidance such as GPS (Global Positioning System), commercial, military, and general aviation (on a much smaller scale) have been replacing traditional paper aeronautical charts with sophisticated moving map displays. Two of the largest motivations for such development are the cost savings of eliminating the paper charts and the

ease of updating electronically stored information compared with paper charts. In light of this, recent studies have compared the use of electronic and paper maps in instrument approach procedure (IAP) charts. In two of the studies, objective data yielded inconsistent differences between the two (Mykityshyn, Kuchar, and Hansman, 1994; Hannon, 1994). However, Hofer, Palen, Higman, Infield & Possolo (1992) and Hofer (1993) found faster information retrieval times for decluttered electronic charts (IAP's) than for paper ones, and in Hofer & Wickens (1997) pilots in a full mission approach simulation flew more effectively with the electronic map than with its paper counterpart. Additionally, in all of the studies, pilots consistently preferred a north-up electronic moving map to a traditional paper map.

As with paper aeronautical charts, most operational electronic moving map displays have been limited to a top down view (-90° perspective). With the addition of digitized terrain databases, 3-dimensional renderings of an area are now possible, eliminating the exclusivity of the 90° map perspective. With the available computer technology, one of the more notable features inherent in the design of electronic maps is the customization of features available. Such customization could be pilot selectable and might include overlaying the terrain depiction with aeronautical chart information, adding or deleting information from the chart, selecting fixed vs. rotating map perspectives, 2-D or 3-D renderings, static or dynamic images, and scaling or zooming functions.

With the increasing availability of different presentation options, it may be tempting to integrate every available feature into this type of display so that the FFOV and the electronic map are close in physical identity. However, if there is a requirement for continuous map updating, driven by aircraft motion, then the technological demands on such a system at present are computationally and financially cost prohibitive. The issue confronted by electronic map designers is how to configure maps with available features so that they can best serve the navigational checking task, within the context of technological feasibility. In other words, in light of the technological constraints, how much can the physical identity of the map with the FFOV be degraded and still support efficient navigational checking? The answer to this question should be derivable from a valid information processing model that reveals the source of the cognitive transformations necessary to accomplish this task.

Present Research

Literature that has investigated variables affecting this model has shown that, when comparing two images (FFOV and map in a navigational checking task) and attempting to establish whether or not the scenes or images are congruent, there is a cost in efficiency (response time and accuracy) as the physical difference between the two increases. This degradation in efficiency is directly related to the amount and type of cognitive transformations necessary to bring the two into congruence to determine if, in fact, they are the "same" or "different." This finding has direct implications to the ideas fundamental to the task of air navigation and to the design of electronic maps. Although many effects have been modeled with varying degrees of consistency, azimuth angle effects have been perhaps the most consistent and replicated (Cooper & Shepard, 1973; Cooper & Podgorny, 1976; Metzlar & Shephard, 1974; Shephard & Metzlar, 1971; Yuille and Steiger, 1983; Goldberg, Maceachren, and Korval, 1992; Eley, 1993 Aretz and Wickens, 1991). Data to support similar transformations along other variables within this model are scarce, particularly with regard to elevation angle disparity and image complexity, and feature type. Yet all three of these issues have tremendous impact on the design of electronic maps. Therefore, the present experiment will attempt to reach some consensus as to the effect of these variables, and recommendations to the design of electronic maps will be made as a result.

Methods

Subjects

Twenty-six subjects, ranging from eighteen to thirty-one years of age (mean=20.3), were recruited from two different instrument flight classes taught by the Institute of Aviation at the University of Illinois. All subjects had only a private pilot rating with a range of flight time from sixty to two hundred hours (mean=97.4). All subjects received the same instructions and were paid $5 per hour for their participation. Subjects were paid an additional $1 bonus as an incentive if they achieved at least 90% accuracy. Twelve subjects (46%) were paid this bonus.

Apparatus and Materials

An Evans & Sutherland SPX500 generated the FFOV and projected it onto a 7' by 10' projection screen. A Silicon Graphics IRIS computer with a sixteen inch diagonal monitor was used to present the electronic map, as well as to record response times and accuracies. Subjects sat in a chair approximately 32" from the IRIS computer monitor (approximately 26° visual angle to center of screen) and 11 feet from the projector screen. To make a response, subjects used a computer flight simulator joystick in their right hand, pressing either a button to indicate "different," or squeezing the trigger to indicate "same."

The Task

During each trial, subjects were presented with static images of scenes from a digitized "world" generated by both the Silicon Graphics and the Evans & Sutherland computers. The "world" consisted of a 15 square mile area depicting roads, bridges, towns, mountains, and rivers, as well as, other natural and man-made features and structures. Although both the map and FFOV were computer generated, the Evans & Sutherland display offers digitized scenery that is significantly more realistic and detailed than the Silicon Graphic Iris display. The subject's task was to compare the two computer generated images and determine if they depicted the same geographic location. "Different" alterations were created by the experimenter and were generated by changing the appearance, shape, orientation, or existence of a few features located toward the center of each scene. Differences could be found with roads, bridges, and buildings, or with rivers, mountain ranges, or plateaus. The images may have looked identical except for one or two changes or deletions. The map was always presented in a "track-up" position, yielding no azimuth differences between the map and the world.

Experimental Design

The scenes that were presented to the subjects were varied on three parameters (independent variables) in a completely within subjects design: 1) Elevation angle deviation, 2) Complexity, and 3) Feature type. Each of these variables is detailed below:

1) Elevation angle deviation (sixteen levels) - Six different map elevation angles (15°,30°,45°,60°,75°,90°) and five different FFOV elevation angles (15°,30°,45°,60°,75°) were used. Although there are thirty possible combinations of angular deviations, only sixteen possible sin deviations exist, hence there were sixteen levels of this variable.

2) Complexity (three levels) - Scenes were categorized by number of features available for comparison. Scenes that were considered the lowest level of complexity had relatively few features for comparison. For instance, the scene may have had one road with a bend in it, or a terrain feature with a single peak. A scene defined by the third level of complexity might have had numerous roads with bends in them, several peaks with rivers that bend around them, or an industrial complex with several roads and waterways surrounding it. Scenes depicting the second level of complexity would be somewhere in between.

3) Feature type (two levels) - Scenes were further divided into two levels of feature type. They were categorized as depicting primarily natural features, or primarily cultural (i.e., man-made features).

Procedure

Each subject participated in one session lasting approximately one hour. After reading the instructions, subjects were situated in the chair and verbally briefed salient highlights of the instructions by the experimenter. Subjects were then presented 8 practice trials to familiarize them with the task. The practice trials were followed by 144 experimental trials. At the start of each trial, a small "X" appeared in the center of the IRIS display and subjects were instructed to fixate on it. After three seconds, both the IRIS display (map), and the screen projection of the Evans & Sutherland computer (FFOV) appeared. Subjects then responded either "same" or "different" via the joystick as soon as they made their decision. Subsequently, both screens went blank for two seconds, and a new trial was initiated with the appearance of the "X."

Results

A 16x3x2 (elevation angle deviation x complexity x feature type) analysis of variance was performed using response time and accuracy as separate dependent measures. Response time results indicated significant main effects for all variables. Additionally, there were significant interactions between elevation angle deviation and feature type, and complexity and feature type.

Analysis of variance on accuracy revealed significant main effects with all variables and significant interactions with everything but sin-sin x complexity. A summary of the relevant findings is detailed below.

Elevation Angle Deviation Effects

Figures 1 and 2 present the response time and accuracy data as joint functions of map angle (the abscissa) and FFOV (the parameter). Although the experimental design did not allow a statistical examination of map angle, unconfounded by disparity, visual examination of the data fails to reveal any apparent monotonic trend of performance across map angle (levels on the abscissa), for response time; however, there does appear to be a general downward trend in accuracy as map elevation angle increases. The data also suggest a trend for performance to be poorest at low FFOV angles, although this trend is contraindicated by the data for the lowest map angle (15°). Additionally, the data indicate that the least amount of variance in response times across all FFOV angles occurs with a 45° map angle while variance in accuracy remains fairly constant for all map angles except for 60°.

Figure 1: Response time as a function of the interaction of map and FFOV angle

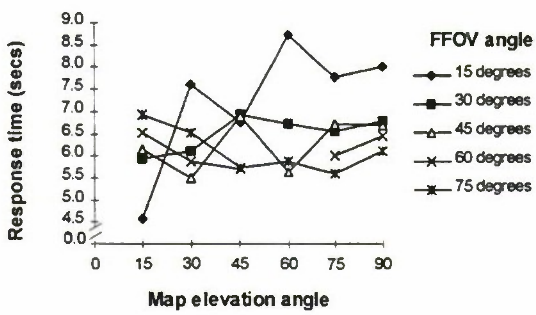
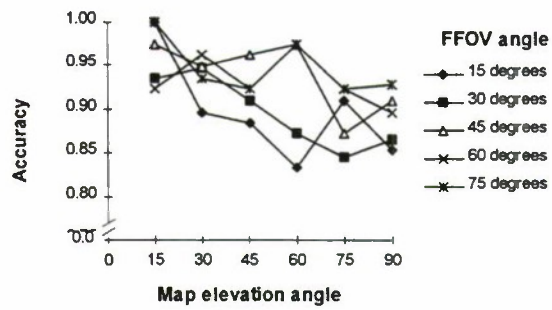


Figure 2: Accuracy as a function of the interaction of map and FFOV angle



Complexity

The analyses of variance reveal a significant effect on response time [F(2,2399)=388.118, p<.001] and accuracy [F(2,2399)=2.631, p=.001] suggesting there were differences in performance affected by complexity level. For accuracy, this effect is monotonic across the 3 complexity levels. However, for response time, it does not distinguish between the highest two levels. Indeed, this may represent a slight speed accuracy tradeoff at high complexity levels; there may be a tendency on some trials to make a "fast guess." As we see below, this guessing strategy may be reflected only with the most difficult natural features.

Feature Type

The ANOVA's also reveal significant differences on both response time [F(1,2399)=25.36, p<.001] and accuracy [(1,2399)=18.239, p<.001] between natural and man-made features. As figures 3 and 4 indicate, scenes that depicted primarily man-made features were responded to approximately 0.7 seconds faster and were 4.9% more accurate than scenes depicting primarily natural features. In a post-experiment questionnaire, subjects were asked "What features were most helpful when making judgments? In other words, what features helped you compare the two views faster?" On average, man-made features were listed 37% more frequently than natural features (48/35).

Figure 3: Response times as a function of feature type

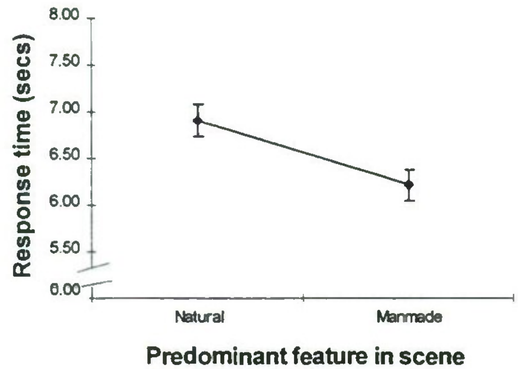
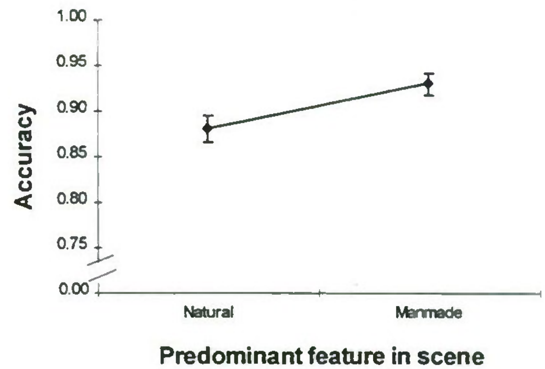


Figure 4: Accuracy as a function of feature type



Interaction

Figures 5 and 6 depict the significant interactions between elevation angle disparity and feature type (response time [$F(15,2399)=1.666$, $p=.051$] and accuracy [$F(15,2399)=2.010$, $p=.012$]). As the deviation increased, it appeared that scenes depicting primarily natural features were more negatively affected than scenes depicting cultural or man-made features.

Figure 5: Feature type interaction with sin of the angular deviations (response time)

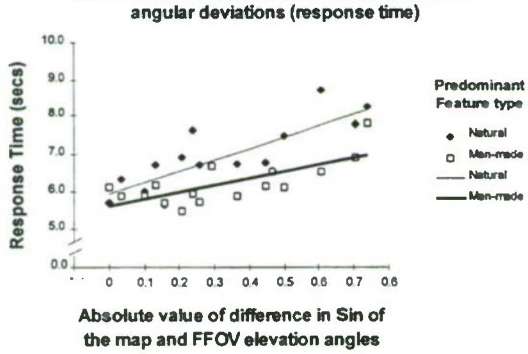
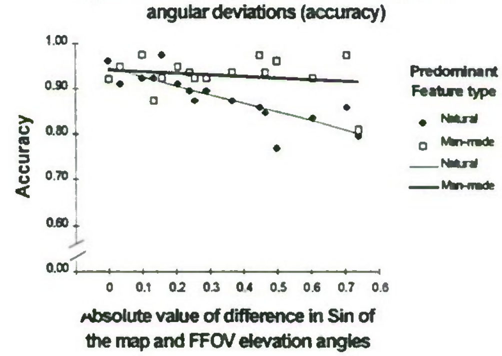


Figure 6: Feature type interaction with sin of the angular deviations (accuracy)



Implications for Map Design

Ultimately, the effects of these variables should be of significant interest to the designers of electronic maps. The general assumption of this paper is that including all possible features on a map may not be cost effective and possibly may not be necessary. The question remains as to which features need to remain in order to maintain efficient navigational checking. The independent variables are discussed in terms of their relevance to map design.

1) Elevation angle disparity. The practical significance of modeling the elevation angle disparity in terms of the difference in the sin values has important applications for the speed of dynamic updating of maps. Consistent with the findings and recommendations of Schreiber et al., if small disparities do not matter (e.g., higher FFOV's coupled with higher map angles) then dynamic updating does not need to be so fast (to keep up with minor deviations). Instead, discrete changes in map angle should be presented more frequently, only when the image differences are fairly large and, therefore, when disparity matters. At FFOV elevation angles of 45° or greater, the efficiency of navigational checking is only slightly affected by changes in the elevation angle of the map. However, it does appear that higher map angles (75°, 90°) do not promote the efficiency that mid-range map angles (30°, -60°) afford. What is definitive in this study is that 90° maps are not the most effective, and it appears that 45° maps offer the least variance in efficiency across all FFOV angles. This finding is consistent with the work of Yeh and Silverstein (1992), who examined an altitude/depth judgment task, and Ellis and Kim (1985) who had subjects perform a tracking task.

It is conceivable that most low-level navigational checking would occur at FFOV angles less than 45°. For example, the pilot of a combat aircraft at an altitude of 500 ft, traveling at 500 knots, will be looking ahead of the aircraft at least 5 miles, producing an FFOV elevation angle of less than 1°. A pilot would have to be looking less than 200 yards in front of the aircraft to have a 45° FFOV, a distance which is highly improbable given the speed and nature of most combat missions. Further analysis of the data in figures 1 and 2 reveals an 8% loss in accuracy for large disparities (e.g., a 90° map with a low angle FFOV) and, a 1.5 second increase in response times for large disparities. Although, this may seem negligible, for a combat aircraft flying at 500 knots, this latency translates into almost 1/4 mile distance traveled, a nontrivial span. A similar scenario can be envisioned for general aviation pilots, however, the range of FFOV angles would be much greater. Most general aviation cross-country or navigation occurs between the altitudes of 1000 to 5000 feet. Even at 5000 feet, a 45° FFOV would center a pilot's attention only 1 mile ahead of the aircraft, a distance traveled in 30 seconds or less. Given the nature of navigation and the range of FFOV angles typically employed, we recommend that, for higher altitude navigation (above 5000 ft), static maps at 45° elevation angle be utilized and efficient navigational checking maintained. Below 5000 ft, a breakdown in efficiency could be expected to occur with static maps, therefore, dynamically updated maps that match the momentary slant angle of the pilot looking out of the aircraft (FFOV) would be recommended.

2) Complexity. Recommendations for map design with respect to complexity would be indeterminate. The general finding of this study is that increasing complexity reduced performance, but the nature of the experimental design prohibited the determination of whether it was map or FFOV complexity that mattered the most. Complexity does effect navigational checking, yet how does one design for it? How does one manipulate the complexity of a map to improve efficiency? A variable of interest that was not manipulated in this study is map simplification. It is possible to remove features of a map and still maintain efficient checking, but the question remains as to which features of a map to remove. Given that the data suggest that judgments of orientation and position using man-made or cultural features are made faster and more viewpoint invariant (Biederman & Gerhardstein, 1993), we recommend highlighting those features to make them more salient and not consider removing them in any map simplification. Given the overall effect of complexity on the navigational checking process, designers should be cognizant that the process will be different for different types of scenarios. Navigating in a desert scenario will be much different than navigating in a mountainous or urban setting. In a similar manner, operators should be aware of these differences as well in that, training for a particular scenario may not be appropriate for other operational possibilities.

Conclusions and Future Research

In this study, we attempted to model the effects of elevation angle disparity, complexity, and feature type within a framework that explains the cost in efficiency of these cognitive transformations on navigational checking. The primary concern of this research was to model the effects of elevation angle disparity of the map and FFOV angles. We also tried to establish the costs imposed by differences in levels of complexity and feature type. Although we established very clear differences in performance caused by manipulation of our variables of interest, further research is necessary. The paradigm used in this experiment was that of a same/different judgment task. As we formulate a model that incorporates the effects of various factors on the efficiency of navigational checking, we feel this paradigm is appropriate. However, "sameness" is the probable end result of most navigational checking iterations, and other paradigms may be employed and could reveal even more dramatic effects of the various manipulations. Real-time simulation using dynamic updating for visual navigation scenarios and geographic orientation decisions, or possibly target acquisition and identification, may yield further evidence to either support or dispute our modeled effects. Also of interest are the effects of differences in the speed of dynamic updating, (e.g., continuous at 60 Hz. or once per second). The issues surrounding map simplification also need to be investigated. How much can the number of shared features between a map and the world that it represents be reduced and still allow pilots to know where they are? The literature reveals other variables of interest relevant to map design such as scale or zooming, color, clutter, or vibration, and more should be done with respect to applying these findings to the design of dynamic electronic, 3-D maps. Given the technological trends in aviation, the study of these factors is both timely and appropriate.

Acknowledgments

This research was funded by contract N000014-93-1-0253 from the Naval Air Warfare Center Weapons Division, China Lake, California. Bob Osgood and Jan Stiff were technical and scientific monitors. The authors acknowledge the valuable programming contributions of Sharon Yeakel.

References

Aretz, A.J. (1991) The design of electronic map displays. Human Factors, 33(1) 85-101.

Aretz, A.J., & Wickens, C.D. (1992). The mental rotation of map displays. HumanPerformance,5(4) 303-328.

Biederman, 1., & Gerhardstein, P.G. (1993). Recognizing depth-rotated objects: evidence and conditions for three-dimensional viewpoint invariance. Journal of Experimental Psychology: Human Perception and Performance, 19(6), 1162-1182

Bond, D.F. (1991, April 1). Army speeds up helicopter enhancements in response to Desert Storm problems. Aviation Week and Space Technology,134, 24-25.

Cooper, L.A., (1980). Recent themes in visual information processing: A selected overview. In R.A. Nickerson (Ed.), Attention and Performance (pp 319-345). Hillsdale, NJ: Erlbaum Associates.

Cooper, L.A. & Podgorny, P. (1976). Mental transformations and visual comparison processes: Effects of complexity and similarity. Journal of ExperimentalPsychology: Human Perception and Performance, 2, 503-514.

Eley, M.G. (1993). The differential susceptibility of topographic map interpretation to influence from training. Applied Cognitive Psychology, 7, 23-42

Goldberg, J.H., Maceachren, A.M., and Korval, X.P. (1992) Mental image transformations in terrain map comparison. Cartographica, 29(2), 46-59.

Hackworth, D.H. (1991, November 18). Killed by their comrades (friendly fire). Newsweek, 118, 45-46.

Hannon, D.J., (1994). Effects of a moving map display on pilot performance on instrument approach procedures. SAE Technical Paper Series presented at Aerotech '94, Los Angeles Oct 3-6

Hofer, E.F., Palen, L.A., Higman, K.N., Infield, S.E., & Possolo, A. (1992). Flight Deck Information Management-Phase II, Final Report to the FAA on Contract DTFA 01-90-C-00055. Boeing Commercial Airplane Group Document No. D6-56305-1.

Hofer, E.F. (1993). Flight Deck Information Management-Phase III, Final Report to the FAA on Contract TFA01-90-C00055.

Mykityshyn, M.G., Kuchar, J.K., & Hansman, R.J. (1994). Experimental study of electronically based instrument approach plates. The International Journal of Aviation Psychology, 4(2), pp. 141-166.

Schreiber, B., Wickens, C.D., Alton, J. & Renner, G. (1996). Navigational Checking: Implications for Electronic Map Design Proceedings of the 40th Annual Meeting of the Human Factors and Ergonomics Society. Santa Monica, CA:

Shepard, R., & Metzlar, J. (1971). Mental rotation of three dimensional objects. Science, 171. 701-703.

Williams, H.P., Hutchinson, S., & Wickens, C.D., (1996). A comparison of methods for promoting geographic knowledge in simulated aircraft navigation. Human Factors, 33, 355-376.

Yuille, J.C. & Steiger, J.H. (1982). Nonholistic processing in mental rotation: some suggestive evidence. Perception and Psychophyiscs, 31, 201-209

This page intentionally blank.

ENHANCED SPATIAL STATE FEEDBACK FOR NIGHT VISION GOGGLE DISPLAYS

E. N. Bachelder
Massachusetts Institute of Technology
Department of Aeronautics & Astronautics
Cambridge, MA

R. John Hansman Jr.
Massachusetts Institute of Technology
Department of Aeronautics & Astronautic
Cambridge, MA

ABSTRACT

A preliminary study was conducted to investigate the use of visual flow cues as an aid to ground and vertical drift awareness during helicopter flight and targeting while using night vision goggles (NVGs). Three displays were compared: 1) *NVG* display: simulated NVG image of cockpit and external environment; 2) *Overlay* display: NVG image with an overlay of a flow cue field and a surrounding wire-frame globe; 3) *Cut-out* display: same as the *Overlay* display but with symbology removed from the central region (leaving an unobscured 20 degree field of view of the NVG image). Three levels of contrast were also compared using each display type. The visual scenery was displayed to subjects using a helmet-mounted virtual reality device that had a 40 X 50 degree field of view liquid crystal display

The study involved six pilots. Three tasks were given: 1) Search task: designate enemy targets with a helmet-mounted sight (no flight control inputs); 2) Hover task: null out all translational and yaw rates while in a hover; 3) Search/Hover task: perform both Search and Hover tasks simultaneously. These tasks were conducted in a fixed-based helicopter simulator which used the dynamics of a small-scale model helicopter. The following performance measures were collected: 1) Pilot ability to detect and recognize targets (Search and Search/Hover tasks); 2) Pilot ability to null translational and yaw rates (Hover and Search/Hover tasks); 3) Time scanning the instrument panel (Hover and Search/Hover tasks). Subjects also rated displays for efficacy in completing the three tasks.

Target detection scores conducted during the Search and Search/Hover tasks were highest using the NVG display, followed by the *Cut-out* display. Root-mean-square (RMS) drift rate error was comparable for all display types in the Hover and Search/Hover tasks, however RMS control input activity in all the translational axes was significantly higher in both rate-cueing displays (*Overlay* and *Cut-out*) than with the *NVG* display. From the control input and drift rate time histories it appeared that the motion cues were more compelling in the *Overlay* and *Cut-out* displays than those perceived in the *NVG* display. A significant decrease in instrument-scanning time (both Hover and Search/Hover tasks) was observed for both the *Overlay* and *Cut-out* displays compared to the *NVG* display, with pilots flying essentially head-out-of-cockpit while using the rate-cueing displays. Contrast was not observed to have a significant effect on hover performance in any of the displays.

1. INTRODUCTION

Rotorcraft night operation in close ground proximity is a difficult problem which requires two concurrent tasks: object recognition (for mission tasks and obstacle avoidance), and vehicle orientation. The motivation for this paper was based on studies of night vision goggle (NVG)-related helicopter mishaps, from which four major areas of perceptual difficulty were identified. These were: altitude, sink rate, horizontal ground drift, and attitude. This evidence, coupled with anecdotal reports, implies that the visual feedback mechanisms which impart perception of these states during day unaided vision can be unreliable or absent during NVG viewing. Attempts to improve the field of view (FOV) and enhance image resolution may improve the situation, but perceptual problems with image-sensing systems are likely to remain areas of concern.

The approaches adopted for addressing these perceptual issues were: 1) Identify the functional requirements to provide visual feedback in a rotorcraft helmet-mounted display (HMD), including the perceptual requirements to support flight control feedback loops; 2) Identify methods to provide required feedback which are consistent with current understanding of visual perception; 3) Develop prototype HMDs which employ the identified feedback mechanisms; 4) Preliminary testing and evaluation of display options using potential helicopter mission profiles.

Based on current work including the NVG mishap studies and several prior surveys,[1,2,3] three major areas have been identified as contributing to NVG visual feedback problems: 1) Loss of peripheral or off-axis visual cues due to narrow FOV in current HMD's; 2) Sensor resolution and noise causing degradation of spatial cues; 3) Motion ambiguity created by head-tracked motion of the sensor.

From the functional requirements analysis, visual flow cueing has been identified as a major area of opportunity for improving performance. Anderson and Dyre[4] conducted an experiment in which subjects were given a circular 15 degree FOV and immersed in a three-dimensional field of dots. The subjects were then moved through this field (visually simulated) and their body sway measured to observe the sensation of self-motion. The results showed that central vision can be stimulated to produce strong sensations of motion both parallel and normal to the viewer's line of sight (LOS). As a powerful tool for stimulating motion perception within and near central vision, flow cues can be a method of conveying rate feedback. It is proposed that presenting these cues on the image periphery will: 1) Leave the primary visual task largely unobstructed; 2) Provide compelling and useable rate-cueing for tasks requiring spatial awareness.

2. EXPERIMENTAL DESIGN

2.1 Overview

This experiment investigated the effects of three different display designs and three contrast levels on target detection and hover performance. The baseline display simulated a night vision goggle (NVG) image as seen from the pilot's perspective. The *Overlay* display superimposed three-dimensional (3-D) symbology on an NVG image to cue vehicle rate to the pilot, and the *Cut-out* display combined the 3-D rate cues with a decluttering strategy while viewing an NVG image. Subjects conducted these tasks in a fixed-based helicopter simulator, with visual scenery displayed to subjects using a helmet-mounted virtual reality device that had a 40 X 50 degree FOV liquid crystal display. The cockpit airframe seen from the pilot's perspective matched the dimensions of the AH-1 Cobra helicopter. Head motion was tracked by an ultra-sound system with a resolution of approximately 0.1 degree. Subjective ratings for each display type were also collected.

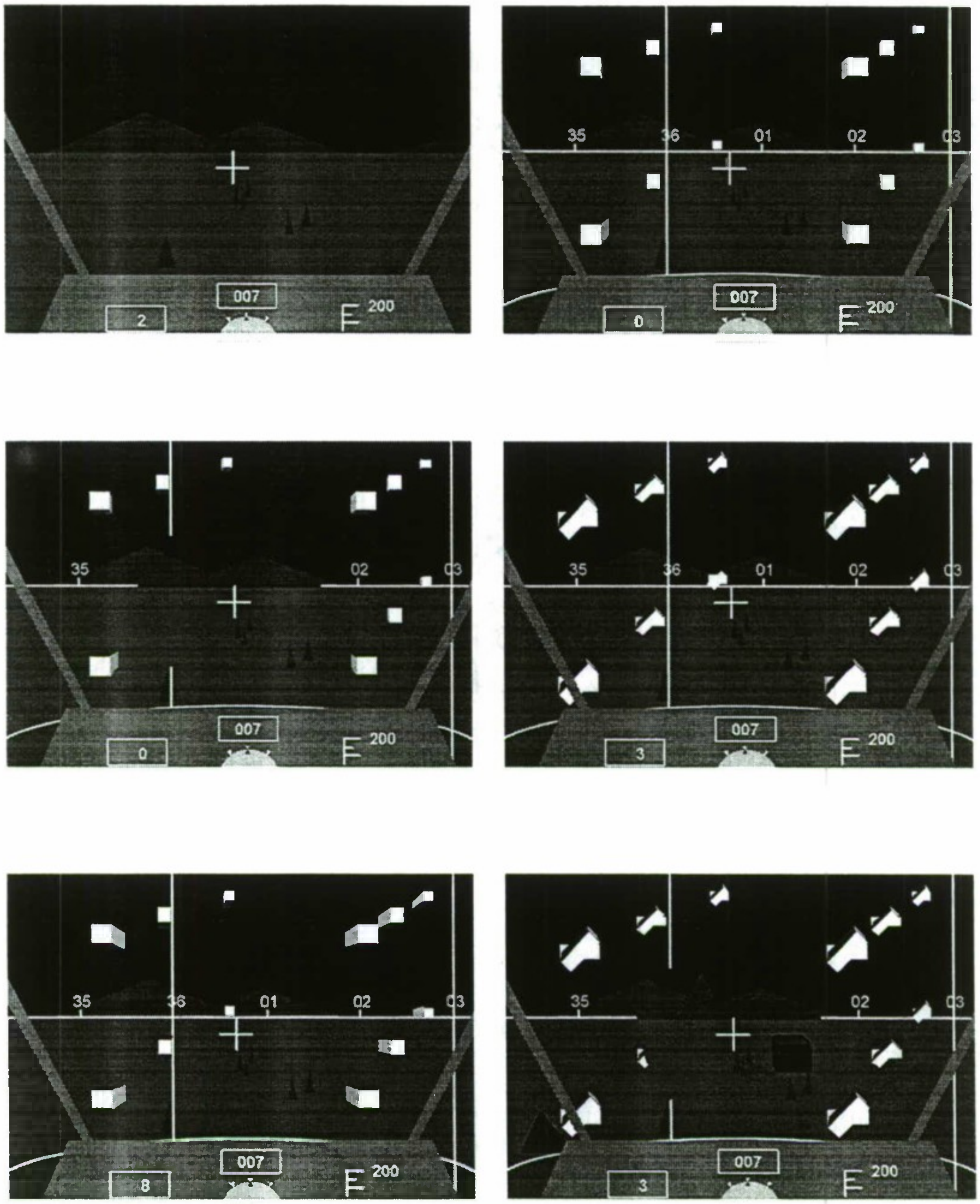
2.2 Display Design

2.2.1 *NVG* display

The *NVG* display is shown in Fig. 1. The pilot is looking out the forward cockpit with the head pitched down 10 degrees, so that the upper portion of the instrument panel is visible. Panel instruments included aircraft heading, radar altimeter, attitude gyro, and total horizontal ground speed. The NVG image (mountains, terrain, trees, airframe, and targets) was rendered monochromatic green, with object intensity mapped from its approximated wavelength.[5] A pair of targeting crosshairs was fixed to the screen's center (line of sight) for all displays.

2.2.2 *Overlay* display

The *Overlay* display in Fig. 2 shows a 3-D lattice of cubes (surfaces shaded blue) surrounding the helicopter. Due to computational constraints, the lattice extended three layers beyond the helicopter in any direction. The lattice was positionally fixed relative to the ground, so that helicopter motion was sensed by a flow of cubes in the opposite direction.



As the helicopter overran a plane of cubes along a given axis, a third distant plane of cubes would appear. As the helicopter built up translational speed, the cubes elongated parallel to the vehicle flight path, so that during translation each elongated block (its center of gravity (c.g.) fixed inertially) would rotate in the direction of the flight path vector. Fig. 2 also shows members of a wire-frame globe (colored white) which was centered on the helicopter c.g. and surrounded the vehicle. The globe had lines of longitude at 30 degree increments of azimuth, and lines of latitude at 20 degree increments of elevation. Azimuth is numerically depicted at 10 degree intervals along the horizon.

2.2.3 *Cut-out* display

Fig. 3 shows the *Cut-out* display, which is identical to the *Overlay* display but with the central symbology removed (except for the targeting crosshairs, which were required for the experiment). As the pilot's head attitude changed, overlaid symbology entering the blanking zone disappeared, and (blanked) symbology exiting the zone reappeared. In Fig. 4, the *Overlay* display is seen during descending and left vehicle translation - the static picture can indicate either left and downward or right and upward motion, but viewing the cues in motion removes this ambiguity. Fig. 5 depicts forward helicopter motion seen using the *Overlay* display. Fig. 6 again shows left and descending motion seen with the *Cut-out* display (note the absence of central cues). Targets (part of the NVG imagery, hence monochromatic green) are shown in Fig. 6.

2.3 Contrast

Three contrast levels were used with each display type: 1) High: nominal object color intensities; 2) Medium: all intensities reduced to 60% of nominal intensities (NVG image, cube lattice, globe); 3) Low: all intensities reduced to 40% of nominal.

2.4 Tasks

2.4.1 Search Task

Ten targets appeared in the external scene when the subject depressed a cyclic button. Five were intended targets shaped as cubes (green since part of the NVG image), and five were distracter targets shaped as pyramids. The objective was to maximize the capture of intended targets. Target capture was performed by head-pointing the display crosshairs over a target and depressing a cyclic trigger switch, subsequently the target would disappear. Every ten seconds a new set of random targets would replace the previous set. Targets appeared within $\pm$ 60 degrees of the aircraft heading in nominal increments of 30 degrees, and actual target azimuth varied randomly $\pm$ 5 degrees from the nominal. Target elevation was randomly generated from -15 degrees (to avoid airframe obscuration of target) to +40 degrees relative to the horizon. Once targets appeared, they remained fixed relative to the ground. A run lasted 60 seconds (six target sets), during which the helicopter was fixed in space at an altitude of 50 feet, requiring no control inputs from the pilot.

2.4.2 Hover Task

The helicopter was automatically positioned and trimmed for a steady-state 50-foot hover. Upon depressing a cyclic button, flight control was transferred to the subject and turbulence initiated. The turbulence model was Gaussian, zero-mean, with maximum gusts of $\pm$ 1 ft/sec in the north, east and vertical directions. Task objective was to null out all translational and yaw rates that developed using any of the available cues (terrain features, panel instruments, artificial flow cues). A run finished after 60 seconds of flight.

2.4.3 Search/Hover Task

The subject conducted both the Search and Hover tasks simultaneously, with the primary objective being helicopter control and stabilization. Each run was 60 seconds in duration.

2.5 Protocol and Data Analysis

All references to significance refer to a 95% confidence level ($p < .05$). The order of tasks each subject received was Search, Hover and Search/Hover. Prior to each task, the subject was briefed on scenarios and objectives. A warm-up session of three target sets was given using each display/contrast combination (9 combinations). For the Hover task, pilots practiced flying with each display until they felt

comfortable with both the vehicle dynamics and displays. Familiarization times ranged from 20 minutes to 45 minutes. A similar warm-up period was given for the Search/Hover task. Following each warm-up, nine 'live' runs were recorded, one for each display/contrast combination (orders were counterbalanced), and subject evaluations collected.

The Analytic Hierarchy Process[6] (AHP) was used to obtain subjective preferences for the three displays for each task. AHP breaks up multiple options into a series of paired comparisons which are then recombined to produce an overall weighted ranking. The scale given to subjects is a measure of dominance of one alternative over the other, shown in Fig. 7.

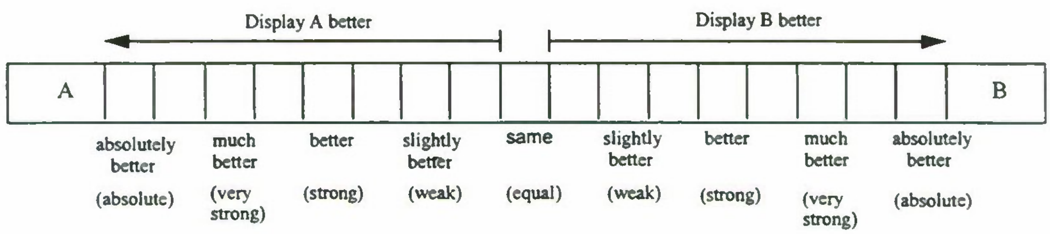


Fig. 7. Dominance Scale Used for Paired Comparison of Two Display Options

2.6 Subjects

Six subjects participated in the experiment. All had actual or simulator helicopter flight experience, ranging from 0-1900 hours actual (mean 950), and 0-75 hours simulator (mean 17). Three subjects were helicopter instructors; all but one had commercial pilot ratings, with the remaining holding a private pilot license. None had flight experience using an HMD. Subject ages ranged from 25 to 42, and all were male.

3. RESULTS

3.1 Target detection

Target detection results are presented in Fig. 8. For both the Search and Search/Hover tasks the NVG display had higher detection scores than the other displays. Search/Hover task scores were significantly lower compared to the Search task scores for all displays.

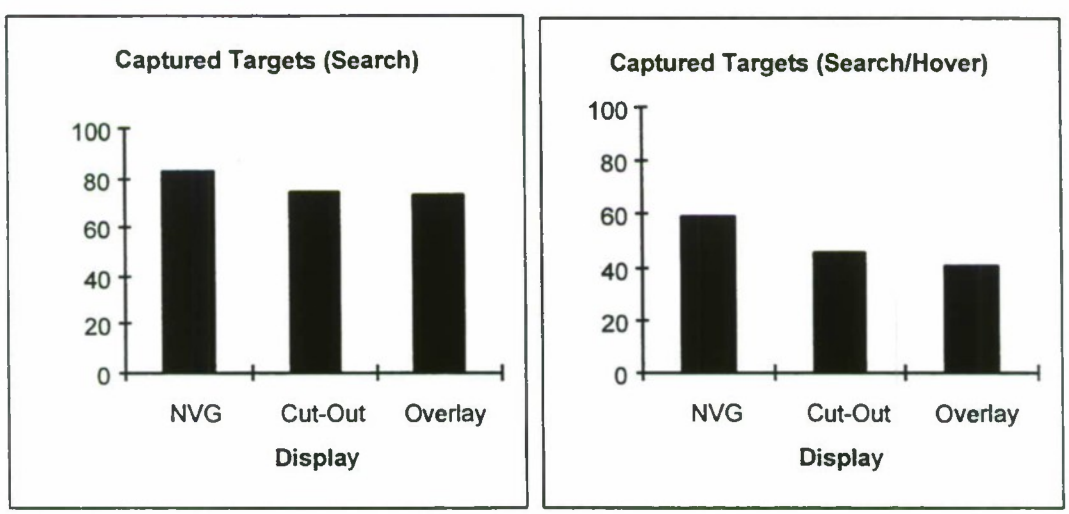


Fig. 8 Target Detection Scores for Search and Search/Hover Tasks

Fig. 9 shows the effect of contrast on target detection. For the *NVG* and *Cut-out* displays, scores were essentially the same going from the High to Medium contrast. At Low contrast, both the *Cut-out* and *Overlay* displays yielded similar scores within each task.

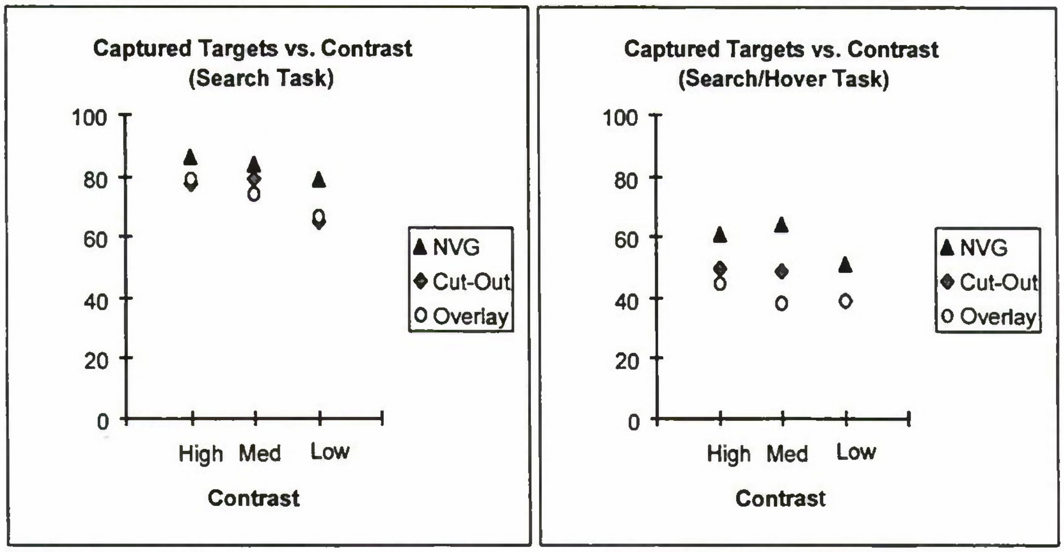


Fig. 9 Target Detection Scores vs. Contrast for Search and Search/Hover Tasks

Fig. 10 shows detection performance for three zones of detection: 0 deg ($\pm$ 15 degrees off the nose), 30 deg (15 to 45 degrees off the nose, both directions), and 60 deg (greater than 45 degrees off the nose, both directions). It is seen that the addition of the high workload Hover task to the Search task results in decreasing detection scores as off-axis angle increases.

Search Task **Search/Hover Task**

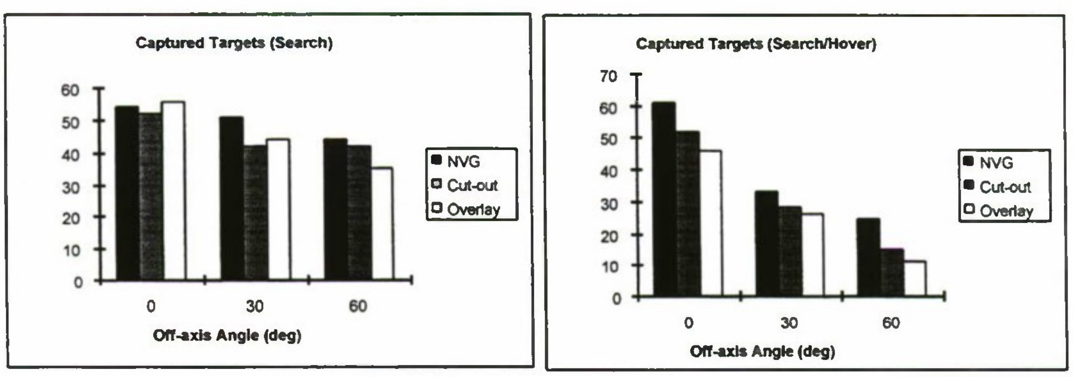


Fig. 10 Off-Axis Target Acquisition Results for Search and Search/Hover Tasks

3.2 Hover performance

Fig. 11 shows typical plots for the *NVG* and *Overlay* displays for drift rate and control input for the three axes of translation (Hover task). In general, lateral and vertical motion appeared to be less oscillatory while using the *NVG* display than with the other two rate-cueing displays. The *NVG* display also resulted in significantly lower control activity in the longitudinal and lateral axes for both the Hover and Search/Hover tasks. Collective activity was also lower with the *NVG* display than with the rate-cueing displays.

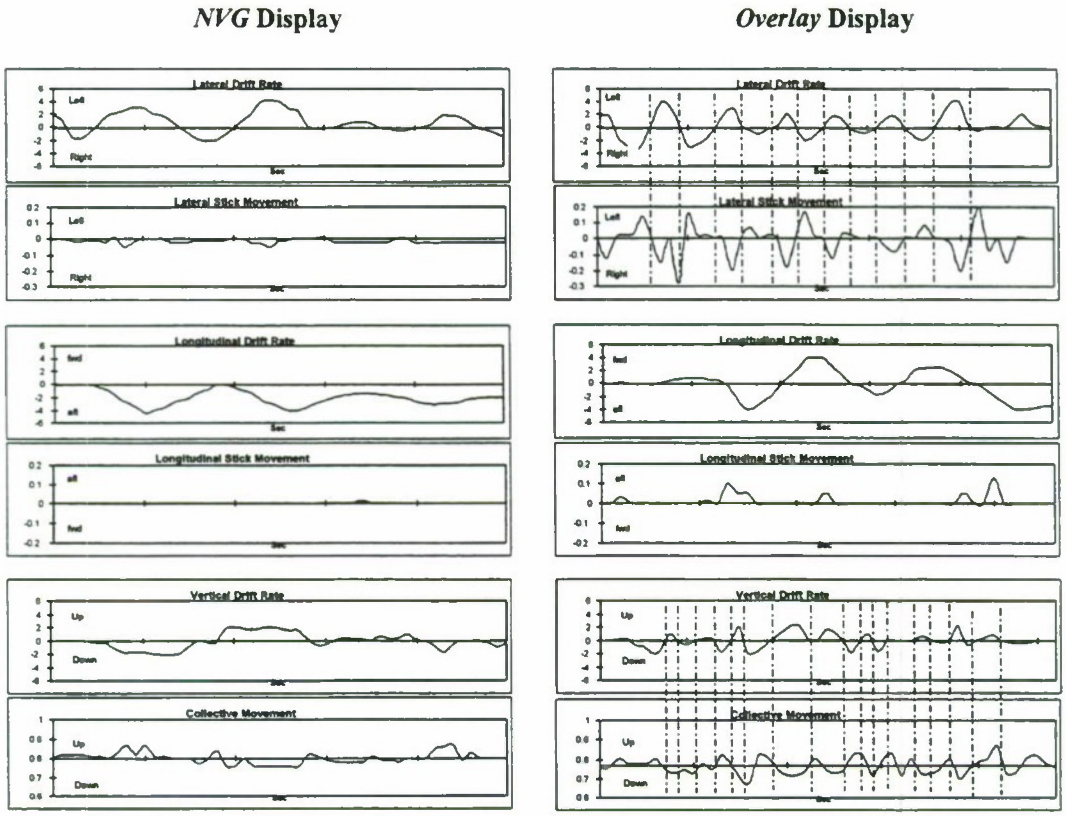


Fig. 11 Case Example of Hover Task Drift Rates and Control Inputs for *NVG* and *Overlay* Displays, Same Pilot (60 sec run)

RMS control input rates are shown in Fig. 12. From this data it appears that the rate cues of both the *Overlay* and *Cut-out* displays imparted stronger motion sensation than occurred using the *NVG* display. RMS drift rates, shown in Fig. 13, were comparable for all three displays for the Hover and Search/Hover tasks.

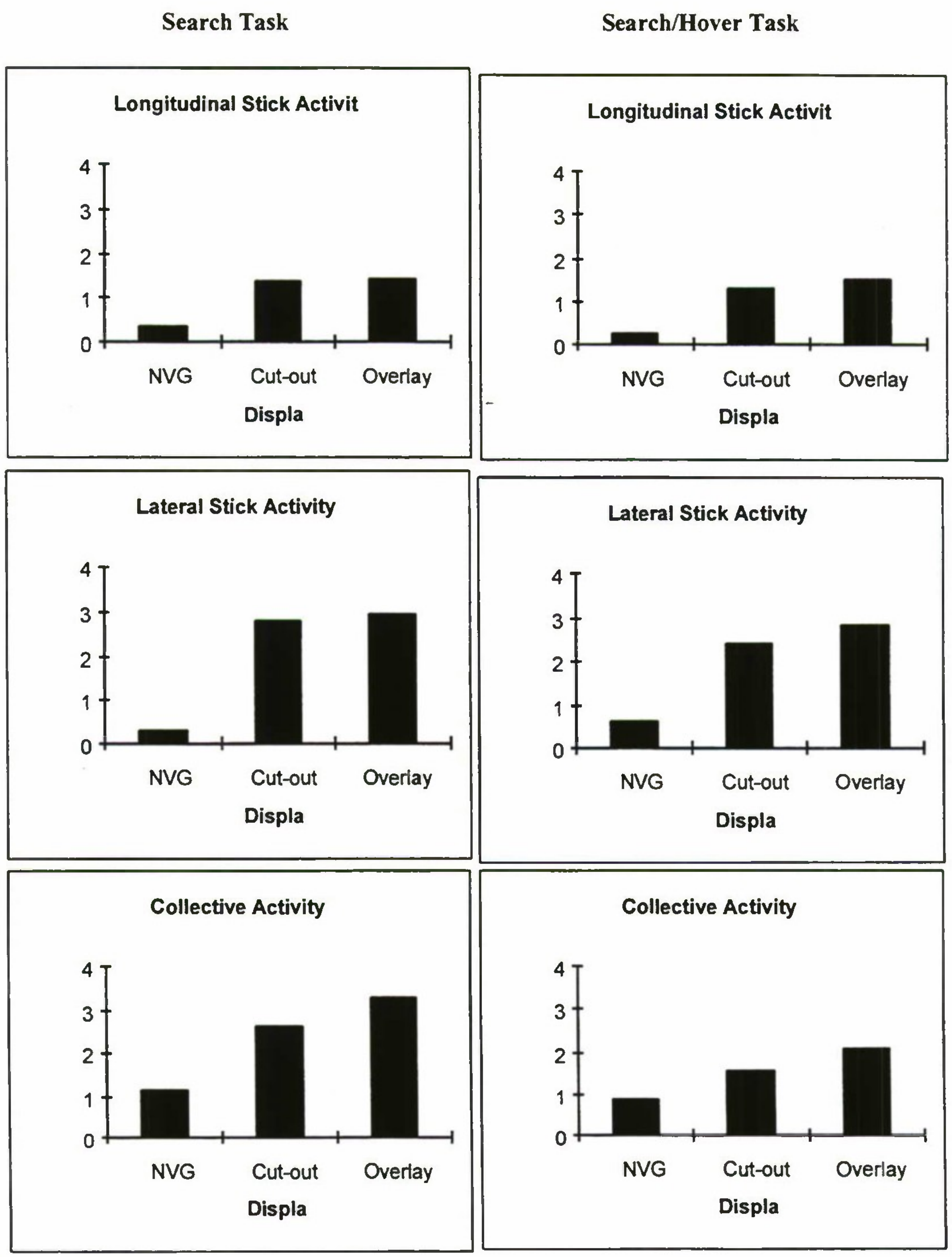


Fig. 12 Control Input Activity for Search and Search/Hover Tasks

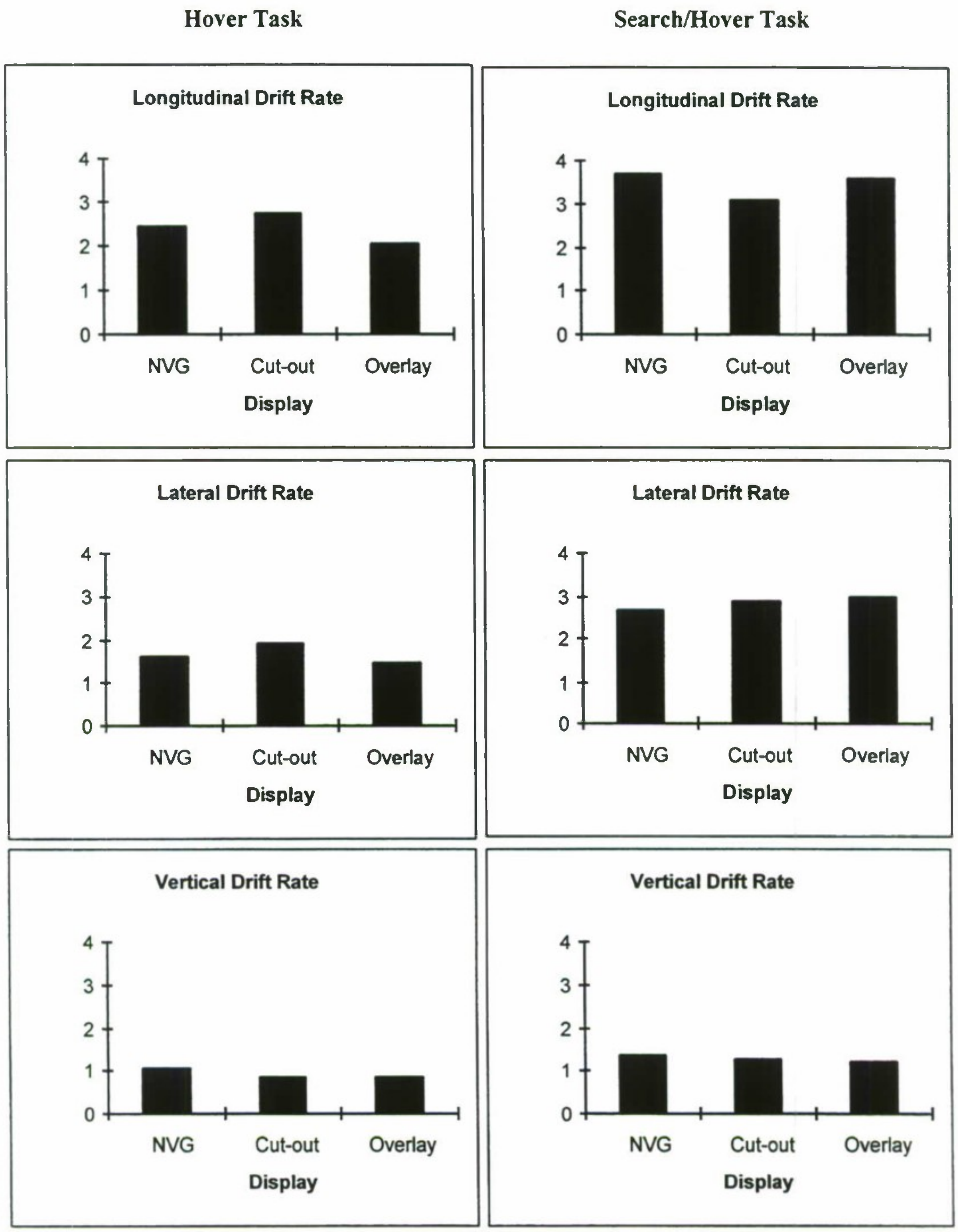


Fig. 13 Hover and Search/Hover Task Drift Rates

A significant decrease in head-down time (Fig. 14) was observed for both the *Overlay* and *Cut-out* displays compared to the *NVG* display (Hover and Search/Hover tasks). Head-down time using the NVG display decreased going from the Hover to Search/Hover task, but there was little difference between the tasks for the two rate-cueing displays.

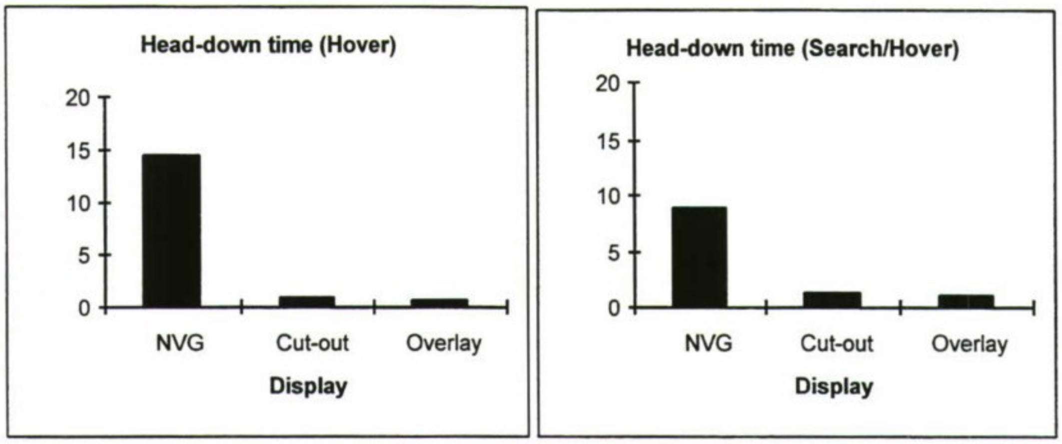


Fig. 14 Mean Head-Down Time for Hover and Search/Hover Tasks (60 sec flight time)

3.3 Subjective rankings of displays

The AHP results are shown in Fig. 15. For the Search task, the *NVG* display was preferred over the other two displays, which agrees with the better detection performance it produced. For the Hover task, the *Overlay* was preferred, and for the Search/Hover task, the *Cut-out* display was preferred.

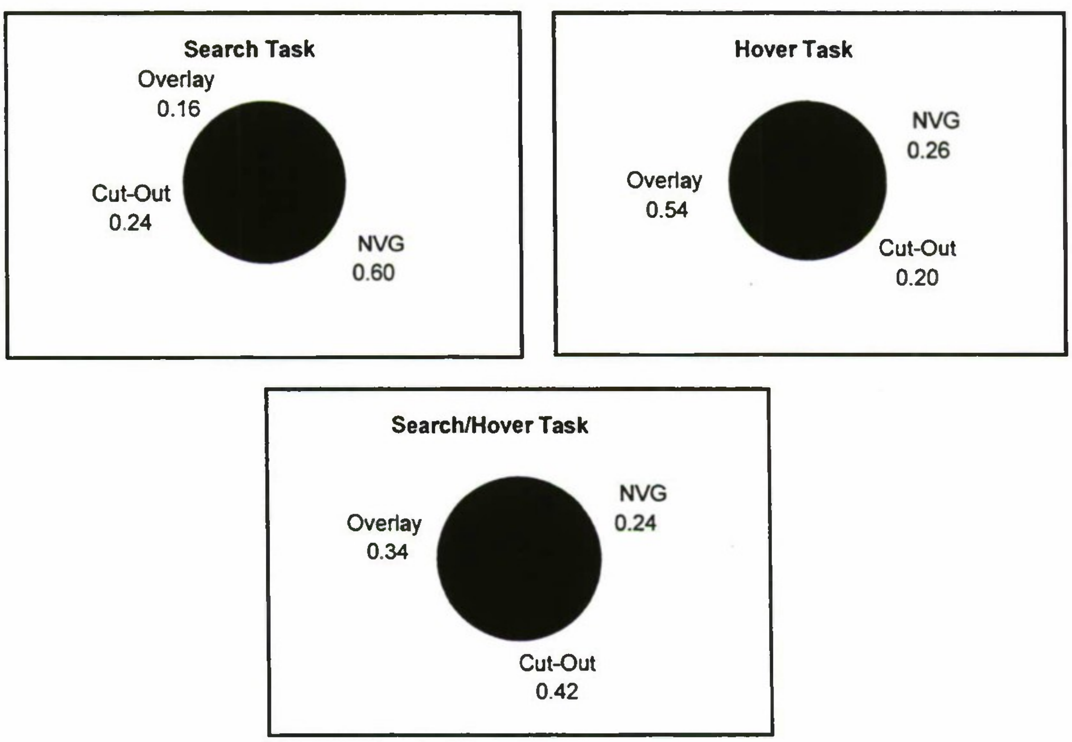


Fig. 15. Weighted Ranking of Displays Using AHP

4. CONCLUSIONS

4.1 Summary of results

- All three displays produced similar drift rate errors in the three translational axes, but control input activity was higher with the rate-cueing displays. From the control input and drift rate time histories, it

appears that the motion cues were more compelling in the *Overlay* and *Cut-out* displays than those perceived in the *NVG* display. A significant decrease in head-down time was observed for both the *Overlay* and *Cut-out* displays compared to the *NVG* display (Hover and Search/Hover tasks). Contrast was not observed to have a significant effect on hover performance in any of the displays.

- Target detection scores were highest with the *NVG* display, followed by the *Cut-out* display. Introduction of the Hover task to the Search task resulted in a significant decrease in target detection scores, this effect being more marked with the rate-cueing displays . Target detection scores during the Search/Hover task were progressively lower with increasing off-axis target angle for all displays.

- Subjective scores indicated a preference for the *NVG* display with the Search task, a preference for the *Overlay* display with the Hover task, and a preference for the *Cut-out* display with the Search/Hover task.

4.2 Remarks on results

The hover performance data showed comparable RMS drift rates, but inspection of the control inputs histories revealed that subjects were more sensitive to drift rates while using the rate-cueing displays than with the *NVG* display, notably in the vertical and lateral axes. The lateral rate behavior using the *Overlay* and *Cut-out* displays is consistent with a control strategy that responds primarily to rate. Bank angle serves as a predictor of lateral rate, and degradation of bank angle sensing in the presence of strong rate feedback could produce results similar to what were observed in this experiment. Seen with a reduced FOV, the Cobra cockpit offers very few markers for cockpit reference, and most of the subjects commented that there was an uncomfortable lack of visible airframe reference while looking out of the cockpit. It is possible that improved cockpit reference cues could reduce some of the drift rate oscillation and high stick activity observed with the rate-cueing displays, and it is also possible that this hover performance was, to some extent, an artifact of the cueing design itself.

For the Search task, the target detection scores for all three displays were comparable. During the Search/Hover task, subjects were essentially flying head-out-of-cockpit with the rate-cueing displays, so one might have expected their detection scores to be comparable to, if not higher than, the *NVG* display's scores for the same task. A possible explanation of why this did not occur is that the artificial rate cues were so compelling that they diminished the effect of and attention given to other visual stimuli. Removal of the *Overlay* display's central flow cues did not appear to effect hover performance, although the subjective scores for the Hover task indicated preference for having the entire cue field in sight. The target detection scores indicated a benefit to having a central declutter zone when using the rate-cues. Subjective ratings favored the *Cut-out* display when conducting the Search/Hover task, indicating that, of the three displays, subjects thought it best balanced the trade-off between external object and drift rate awareness.

5. ACKNOWLEDGMENTS

This study was supported by the Charles Stark Draper Laboratory under grant DL-H-496004. Gratitude is given to the pilots who volunteered their time and effort necessary to participate in this experiment.

6. REFERENCES

1. J. Crowley, C. Rash, R. Stephens. "Visual illusions and other effects with night vision devices," *SPIE*, Vol. 1695, Helmet Mounted Displays III, 1992.
2. E. Bachelder and R. J. Hansman. "Issues in simultaneous HMD display of multi-reference frames for helicopter applications," *SPIE Proceedings: Enhanced and Synthetic Vision*; Paper No. 2736-23.
3. J. Morey and D. Leedom. "Superimposing instrument symbology on a night vision goggle display during simulated contour flight," *Dynamics Research Corporation*, Contract #MDA-903-92-D-0025, 1993.
4. G. Anderson and B Dyre. "Spatial orientation from optic flow in the central visual field," *Perception and Psychophysics*, Vol. 35, 453-458 1989.
5. R. Miller and T. Tredici. "Night vision manual for the flight surgeon," *Ophthalmology Branch, Armstrong Laboratory*, Report AL-SR-1992-0002, 1992.
6. L. Yang and R. J. Hansman. "Application of the analytical hierarchy process for making subjective comparisons between multiple automation/display options," *6th IFAC/IFOR/IEA Symposium*, 1995.

STATISTICAL NETWORK APPLICATIONS OF DECISION AIDING FOR C41 CHALLENGES

Keith C. Drake, Ph.D
AbTech Corporation
Charlottesville, Virginia

Lana B. Atwell
Naval Surface Warfare Center – Dahlgren Division
Dahlgren, Virginia

Abstract

We present an overview of the Command, Control, Communications, Computers, and Intelligence (C4I) algorithm development challenges faced by the Advanced Processors (AP) section of the Naval Surface Warfare Center, Dahlgren Division's (NSWCDD) Munitions Branch. We also discuss CADET – the *C4I Algorithm DEvelopment Tool* – which is poised to address many of these challenges. The NSWCDD is currently pursuing several thrusts in developing and applying new technologies to support the Navy's expeditionary warfare efforts. CADET is a tool which supports many of the algorithm development challenges faced by NSWCDD.

CADET is an advanced "data mining" tool – it automatically finds complex relationships in many types of C4I data, and produces algorithmic models of those relationships. CADET provides a complete environment for importing raw C4I data from many types of sources; data pre-processing and feature selection; producing models that relate a set of input variables to an output parameter; evaluating, validating and implementing models; and performing other types of data and model analysis tasks. CADET's Expert Mining Strategies™ allow intelligent automation of many components of the data mining process. CADET is a significant advance in automated algorithm development software tools, and a substantial improvement of the state-of-the-art for data mining. CADET has also dramatically reduced the time and costs associated with data mining, modeling, and analysis.

This paper reviews current NSWCDD expeditionary warfare thrusts, the fundamental Statistical Network™ data mining technology upon which CADET is based, C4I algorithm development requirements, and example CADET applications.

Acknowledgment

CADET was partially developed under contract to the Naval Surface Warfare Center, Dahlgren Division. Its development under NSWC sponsorship is ongoing.

1. Introduction

The NSWCDD is currently pursuing several thrusts in support of the Navy's expeditionary warfare efforts. These include both algorithm and hardware development for the amphibious-warfare and naval-surface fire-support tasks of automatic target recognition (ATR), modeling and simulation, and data visualization. Specific thrusts include:

- Advanced Processors for Weapons Sensor Fusion (APWSF),
- Advanced Systems for Air Defense (ASAD), and the
- Technology Evaluation Assessment Modeling and Simulation (TEAMS) facility.

One of the primary goals of these efforts is to provide expeditionary warfare forces with decision aiding algorithms. The key to developing useful decision aiding software is effective data analysis and modeling. Often called "data mining," automated algorithms that learn relationships from example data can be used for many data modeling and assessment tasks. The need for robust and automated data mining tools is immense.

Analytical development of data processing models requires specification of relevant relationships describing all aspects of the behavior of the system, often referred to as modeling from "first principals." While this approach can produce accurate data assessment algorithms for real-time C4I applications, it typically presents several challenges. These primarily result from the complexity and uncertainty associated with the majority of decision aiding scenarios, in which case the exact relationships required for decision aiding are not known. In these cases an empirical approach, where decision aiding algorithms are produced by an automated computer system that "mines" relationships from example data, is required.

While many approaches exist for developing tactical and strategic decision aiding algorithms, the key consideration is not whether a system meeting performance requirements *can* be developed but rather the *time and cost* associated with developing a sufficient system. CADET meets this goal by integrating Statistical Network data mining, intelligent heuristics, and automation. The goal of CADET is to provide a streamlined data mining environment to greatly assist in prototyping, validating, and implementing a wide range of software algorithms in support of the TEAMS facility C4I objectives [1].

2. Current NSWCDD Expeditionary Warfare Thrusts

To prepare the battle space and successfully accomplish their missions, Naval Expeditionary Forces (NEF's) rely on an integrated, yet disparate collection of sensors deployed by various means throughout the amphibious operating area. However, through the Science and Technology (S&T) Round Tables for Expeditionary Warfare and Littoral Warfare, the Belisarius series of Workshops and Wargames, and the Littoral Operations 2020 Game Series, wide-raging deficiencies have been identified in current sensor capability.

Innovative technologies for weapon and targeting sensors, particularly integrating software and processing, are essential. Future weapons will have significantly greater ranges and lethality. Target acquisition must extend beyond the ranges of those weapons. And in a significantly more dynamic and fluid battlefield, greater situational awareness and positive target identification are imperative. Innovative algorithms and advanced processors for specific targeting and weapon system sensors are fundamental to the success of NEF in the future battle space.

The concepts for future NEF operations is presented in FMRP 14-21, *Operational Maneuver from the Sea*, (OMFTS) [2]. This concept was reiterated at the Commandant's Warfighting Laboratory, *Sea Dragon* Brief to Industry, 19 December 1995. Figure 1 depicts the complex NEF "sensor-to-shooter" engagement environment with forces deployed ashore.

The Advanced Processors (AP) section at NSWCDD has contributed an evolving expertise to numerous Navy and U.S. Marine Corps (USMC) tasks involving innovative technologies for weapon and targeting sensors: digital signal processing, embedded processors, and algorithms for automated target recognition and digital signal processing.

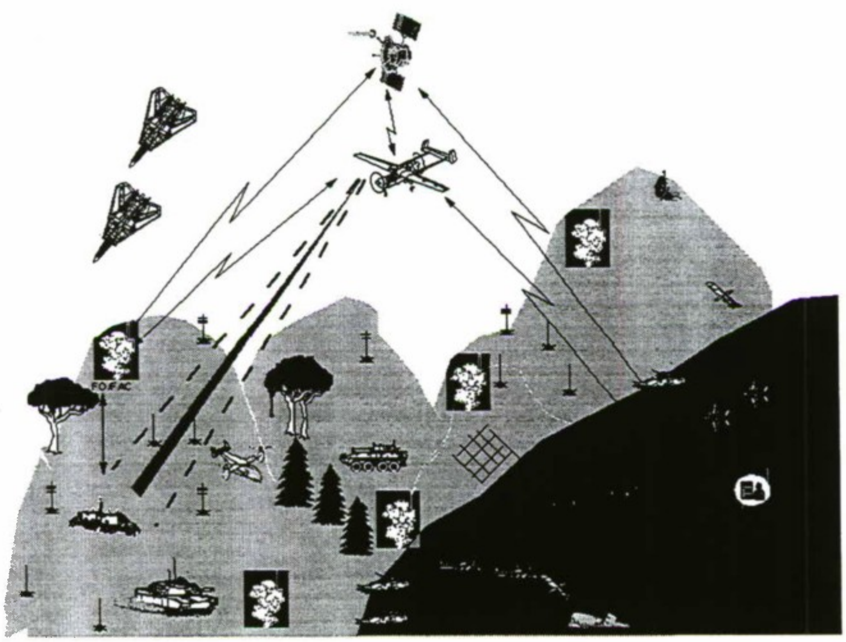


Figure 1 – "Sensor to Shooter" Engagement Environment of Naval Expeditionary Forces

2.1 Advanced Processors for Weapons Sensor Fusion (APWSF)

The USMC's APWSF project grew out of an Independent Exploratory Development (IED) program at NSWCDD. Initiated in October 1988, the program concluded in 1990 with a demonstration of a system that could identify any one of five land-based military platforms from their infrared video imagery [3]. The significance of this effort was twofold: it demonstrated 1) that a system could be implemented with the capability of identifying targets with greater than 83 percent accuracy (from any viewing profile and with degraded data), and 2) that an artificial neural system (ANS) could be implemented with conventional hardware techniques and integrated into a useable target identification system.

Recognizing the potential applications of the technology demonstrated in the IED program, the USMC tasked NSWCDD in 1991 to investigate the application of advanced technologies such as ANS and other neoteric sensor processing technologies to USMC systems under development. Realizing that advanced sensor processing technology would be at the heart of future weapon systems, the USMC created the 6.2 exploratory development task APWSF in FY92 to pursue the development of advanced sensor processing, multiple and disparate sensor fusion, and target detection and recognition [4].

The APWSF 6.2 exploratory development project developed algorithms and system architectures for target classification and identification that have been incorporated in several man-portable or remotely delivered sensor system prototypes.

2.2 Advanced Sensor For Air Defense (ASAD)

Marine Air-Ground Task Forces can deploy anywhere in the world and must maintain freedom of maneuver to position themselves to defeat and/or control the threat. Ground-based air defense (AD) fire units equipped with passive sensors can effectively detect, identify, and engage low- to medium-altitude air threats (at night and during adverse weather) while reducing the probability of detection by enemy forces.

The Avenger and the Man-Portable Air Defense Systems (MANPADS) fire units currently rely on external radar cueing and visual air search, both of which have deficiencies, for target acquisition. Passive sensor technology being explored includes electronic support measures (ESM) and acoustic systems. ESM sensors exploit the radio frequency emissions of aircraft avionics equipment, while acoustic sensors exploit the aircraft's acoustic signature. Both sensors have the ability to provide weapons system cueing as well as non-cooperative target recognition (NCTR). The multisensor integration of these two technologies is being pursued for Avenger and MANPADS applications. Sensors will be mounted on (i.e., organic to) the fire units to provide a stand-alone target acquisition capability. The ability to sense while on-the-move is also being addressed for Avenger.

The principle objective of the ASAD program is to demonstrate technologies to provide a passive target acquisition capability for Avenger (pedestal-mounted STINGER) and MANPADS (shoulder-launched STINGER) fire units. ASAD has demonstrated the ability to passively detect, acquire, and classify fixed-wing, rotary-wing, and unmanned aerial vehicle (UAV) targets in an operational environment within the engagement envelope of shore-based AD (SHORAD) weapons systems. ASAD has reduced technology risks and provided technology alternatives for the demonstration and evaluation phase of the acquisition cycle.

2.3 Technology Evaluation Assessment Modeling and Simulation (TEAMS) Facility

Beginning in FY97, USMC proof-of-concept demonstrations will be conducted within the physical context of an evolving TEAMS facility. Integration will be achieved with emergent NEF modeling and simulation capabilities. Operationally, Marine Light Regiment and Naval Strike Force Assets and capabilities will be supported in the conduct of USMC operations. Measures of effectiveness will be determined at the engagement level.

Tactical and phenomenological realism will enable detailed system performance and military impact assessments to be performed. Material and combat developers who are charged with developing, procuring, and deploying effective sensor-to-shooter systems will have a much needed facility. Expeditionary warfare and USMC concept developers will be able to use the facility to support and execute Advanced Warfighting Experiments to assess OMFTS concepts, doctrines and technology developments. Developers will have a facility that supports combat development, acquisition and engineering processes. Test and evaluation specialists can also use the tool to plan, asses, and augment their processes as well.

Engineers, analysts, and other technology developers are expected to be the principal users of the TEAMS facility.

Operationally, the primary benefit of the TEAMS facility will be the capability to enhance operational effectiveness by ensuring that the correct sensor is in the right place at the right time, and that the best shooter is provided the necessary information for successful engagement. This will, in turn, enable the NEF to successfully accomplish its mission in a timely fashion with significantly fewer resources and lower costs. Based on technology developed and assessed at the TEAMS facility, combat leader and battle staff training in the employment of sensor arrays can be similarly enhanced by providing a more realistic simulated environment.

The TEAMS facility will provide a high fidelity, synthetic, sensor-to-shooter engagement environment for detailed analysis and evaluation. Emphasis is placed on the tactical level, with aggregation to desired operational levels possible. Tactically relevant, high resolution phenomenology of algorithms, sensors, targets, atmospherics, and obscurants can be evaluated, assessed, developed, and simulated with the technical focus on the sensor-to-shooter interaction. Lethality effects can be emulate for engagement fidelity with emphasis on the littoral domain.

Figure 2 depicts the functional and process design of the TEAMS facility. The facility possesses a core of functionality that effectively integrates and correlates a very high fidelity environment, tactical force representation, and network management. The design allows ease of entry to existing and emergent sensor simulators and models.

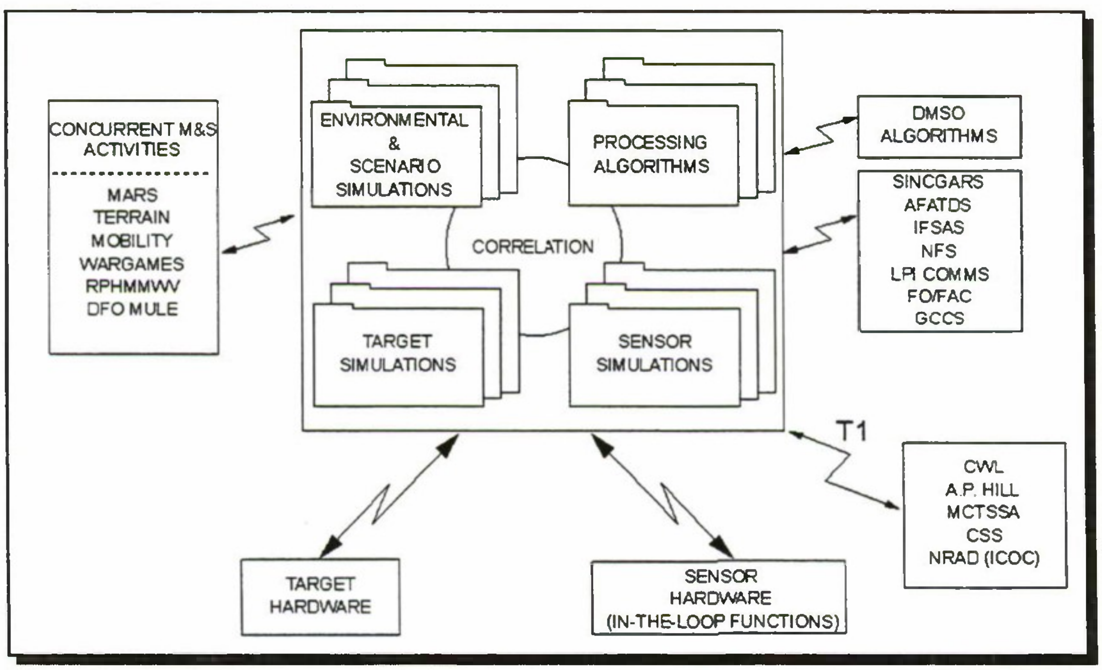


Figure 2 – TEAMS Facility Architecture. TEAMS provides sensor hardware in-the-loop test and evaluation, algorithm development and assessment, simulated target database development, and provides for real-time target input from remote sites.

2.4 NSWCDD Expeditionary Warfare

The exponential growth of technology is transforming the modern battlefield into a high-tech conundrum where the weapons of war perform their missions at speeds beyond the human ability to assimilate information. Compounding the challenge, these capabilities are proliferating as technically advanced nations, driven by increasingly competitive world economics, make even more advanced technologies available to the independent terrorist that were once available only to world superpowers.

Countering such numerous and diverse threats will require weapon and tactical intelligence systems of the future to deal with large amounts of loosely correlated data and to react accurately to rapidly changing information and conditions. Data from multiple, disparate sensors will have to be quickly and concurrently integrated into an accurate representation of the threat environment. The success of future weapons systems hinges on their ability to automatically detect and recognize targets of interest in a chaotic battlefield environment.

The goal of the AP section is to investigate and develop sensor information and processing capabilities that can be applied to new systems requiring the ability to detect and recognize targets in a *real-time* battlefield environment. The technology that has developed within the AP section has broad application to surveillance, scouting, targeting, fire control, information processing, and battlefield management in a variety of military systems. We expect to continue to see advances across the spectrum of science. Through continued innovations, the AP section investigates and develops advanced sensor data-processing technologies to provide future expeditionary-warfare weapons systems with the capabilities of automatic target detection, accurate target classification and, where appropriate, autonomous operation.

3. Statistical Network™ Data Mining Technology

3.1 The Requirement for Automated Data Mining

Analytical development of C4I decision aiding algorithms is an extremely difficult task. For many fundamental and practical reasons, an in-depth understanding of the complex set of interactions between sources of low-level data and high-level decision knowledge required for precise mathematical models typically does not exist. Further complicating this task is determining optimal methods for combining non-sensor, situational and intelligence information with sensor data. Although this analytical expertise is *sometimes* available, it is extremely time intensive (i.e., expensive) to develop robust models which process sensor data and non-sensor information in real-time.

While system behavior can sometimes be described analytically, the robustness of the resulting model is limited when required to operate in a real-world environment. This results from the number and degree of simplifying assumptions required to express the system analytically. Several types of uncertainty require that C4I systems adapt to and reason effectively in a constantly changing battle environment. To effectively deal with the wide ranges of these effects, a methodology for combining highly diverse types of sensor and non-sensor data is required.

The modeling of complex dynamic systems from *examples* of behavior – rather than from a fundamental understanding of the system – is often a more successful strategy. However, most empirical techniques – such as *traditional* neural networks, linear regression, and look-up tables – have basic limitations that restrict their applicability. CADET uses a very powerful and practical empirical technology – Statistical Networks – which have been used to produce accurate *and* real-time models of complex dynamic systems.

3.2 Statistical Network Data Mining

CADET produces relational models (those which relate a set of inputs or observations to a desired parameter estimate) learned inductively from empirical evidence. Relationships which potentially represent a complex process or environment are hypothesized and "scored" according to some criteria which minimizes error. On the basis of the performance of the hypothesized relational model, several refinements and adjustments are made based on a learning mechanism. Traditional statistical regression and neural network approaches offer some utility, but suffer from practical limitations.

Statistical Networks process information with complex mathematical functions. Functions are attractive because they capture a large number of complex relationships in a very compact and rapidly executable form. The CADET statistical learning algorithm produces a network of functional nodes – each node containing a multiple-term polynomial relationship. Polynomial nodes are an extremely powerful method for performing complex reasoning tasks – they are the basis of traditional neural networks and other modeling techniques. They process one, two, or three inputs to compute an output value; and contain a bias or constant term (w_0), and linear, quadratic, cubic, and cross terms. A LINEAR node processes several inputs and contains only the linear and bias terms. The equations for each node type are:

$$\text{SINGLE} = w_0 + w_1x_1 + w_2x_1^2 + w_3x_1^3$$
$$\text{LINEAR} = w_0 + w_1x_1 + w_2x_2 + \ldots + w_nx_n$$
$$\text{DOUBLE} = w_0 + w_1x_1 + w_2x_1^2 + w_3x_1^3 + w_4x_2 + w_5x_2^2 + w_6x_2^3 + w_7x_1x_2 + w_8x_1x_2^2 + w_9x_1^2x_2$$
$$\text{TRIPLE} = w_0 + w_1x_1 + w_2x_2 + w_3x_3 + w_4x_1^2 + w_5x_2^2 + w_6x_3^2 + w_7x_1^3 + w_8x_2^3 + w_9x_3^3 + w_{10}x_1x_2 + w_{11}x_1x_3 + w_{12}x_2x_3 + w_{13}x_1x_2x_3 + w_{14}x_1x_2^2 + w_{15}x_1^2x_2 + w_{16}x_1x_3^2 + w_{17}x_1^2x_3 + w_{18}x_2x_3^2 + w_{19}x_2^2x_3$$

An example Statistical Network is shown in Figure 3. It is a feed-forward network of polynomial nodes processing information from left to right. Each node produces intermediate information which is used as inputs for subsequent nodes. This networking strategy segments the overall relationship being modeled into more manageable components, and simplifies the learning process. Functional networks are synthesized automatically from a flat file database where each column is an input or output parameter (e.g., a variable), and each row contains an example set of the parameters. A hypothesize and test strategy finds the network which best represents the relationships contained in the database.

While individual nodes only allow up to three inputs and are limited to third order terms, employing them in the networking strategy shown in Figure 3 allows the overall network to accept any number of inputs. In addition, because a specific node can contain a third order term, a two-layer network can model a ninth-order relationship. An additional layer allows the modeling of up to 27th order relationships, etc. Therefore, networking relatively simple node types create a powerful knowledge representation.

The Statistical Network learning process produces networks of functional elements which more effectively "learn" complex relationships among features than is often practical with other methods. The key to any machine learning strategy is the learning algorithm itself. It must be able to ***generalize*** from, and not ***memorize***, numerical examples of a problem domain. It must be able to automatically discover relationships to produce a model which performs well for not only training data but independent (i.e., real-world) data. The driving reason for this crucial requirement is ***that all data contains uncertainty***. Noisy, missing, conflicting, and erroneous data are all manifestations of uncertainty in numerical examples.

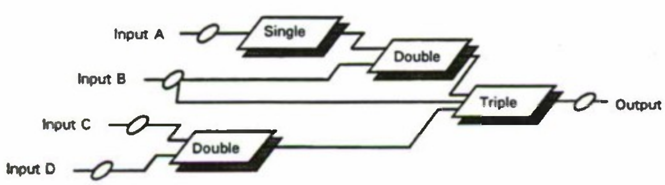


Figure 3 – Example Statistical Network

An effective machine learning algorithm must ***learn relationships*** and avoid ***memorizing noise*** in an ***automated*** manner. Statistical Networks achieve this through the use of intelligent search heuristics to find the optimal network architecture and a modeling criterion to ensure generalization. Following is a top-level summary of the Statistical Network learning algorithm (outlined in Figure 4):

<u>Step 1</u>: Several statistical measures are computed for each database variable such as their mean and standard deviation. The values for each variable are normalized so that they exhibit a mean of zero and a standard deviation of unity, greatly enhancing node regression performed in Step 3.

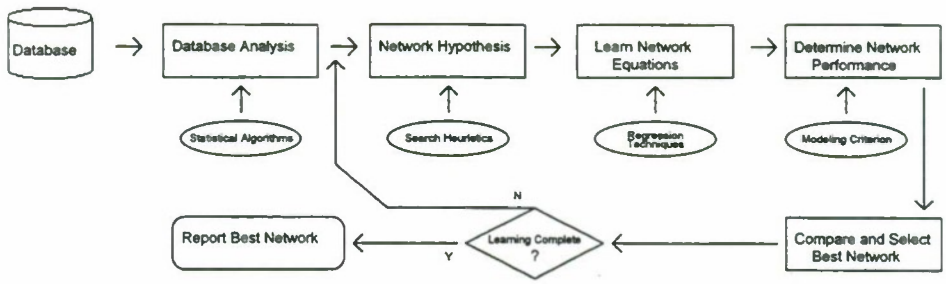


Figure 4 – Statistical Network Learning Algorithm

Step 2: Candidate network architectures are hypothesized using graph-tree network search heuristics. The heuristics employ a survival of the fittest strategy – similar to the underlying concept of genetic algorithms – by hypothesizing more refined versions of networks that have already exhibited promise. Initially, very simple network models are hypothesized (i.e., those which contain only one node). The best of these simple models (as scored by the modeling criterion in Step 4) are then used with the original input parameters as building blocks to hypothesize more complex networks. Search heuristics determine the best manner to combine simpler networks to form more complex ones. This process is repeated (automatically) several times, each providing an additional network layer.

Step 3: For each hypothesized network, coefficients for each node are determined using advanced regression algorithms. The results of Step 3 are values for the coefficients ($w_o, w_1, w_2, \ldots, w_n$) in each network node.

Step 4: Each network is "scored" with the Predicted Squared Error (PSE) modeling criterion, shown in Figure 5. The PSE was developed at Stanford University in the early 1980's specifically as modeling criterion for statistical learning [8]. The network with the best (i.e., least) score is selected as the best for a particular database. The PSE performs a trade-off between network complexity and accuracy to find the simplest network that best models training *and* independent data. It gives an analytic estimate of the network for independent data. The PSE is:

$$PSE = FSE + KP = FSE + CPM\,[\,(2K/N) * s_p^2],$$

where,

- *FSE* is the fitting squared error of the network on the training data.
- *KP* penalizes more complex networks, as they are more likely to overfit training data and, therefore, not perform well on independent data.
- *CPM* is a complexity penalty multiplier, used to vary the emphasis of the *KP* term.
- *K* is the total number of network coefficients in the network model.
- *N* is the number of training observations.
- s_p^2 is an a priori estimate of the optimal model's error variance.

The PSE produces networks which avoid modeling noise and overfitting training data. The Statistical Network synthesis process begins at the left of the PSE curve shown in Figure 5. As the complexity of hypothesized networks increases, the PSE of those networks decrease until the network with the minimum PSE is found. The learning process ends when certain "stopping criteria" are met (Figure 5). These criteria include heuristics which recognize that the learning process is taking place on the upward slope of the PSE curve, and that the best network has been found.

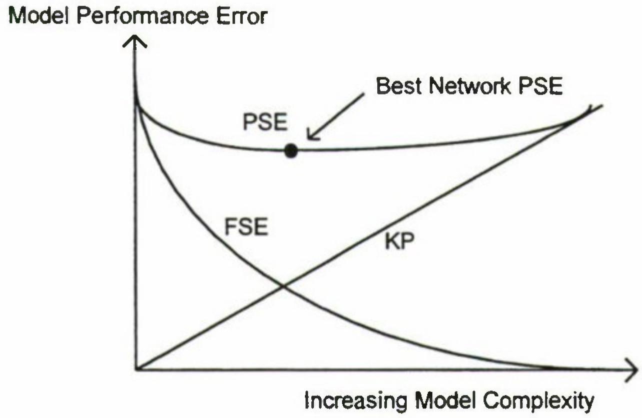


Figure 5 – PSE Modeling Criterion

While Statistical Networks are parametric at the node level, the hypothesis heuristics and modeling criterion at the network level create an *automated* non-parametric process. Therefore, the human user is *not* required to be an integral part of the learning algorithm as *is* required by other approaches. This allows the system developer to focus limited resources on other issues, such as problem definition, system design, model evaluation, and system integration.

Statistical Networks can represent complex relationships and processes that would otherwise require hundreds or thousands of production rules or entries in a decision table. Compared to neural network technology, Statistical Networks excel at estimating continuous parameters, and is much more practical to develop. Because the process is non-parametric, resulting models generally outperform those developed with linear regression. The approach is an empirical one to learning and generalizing relationships and, therefore, does not require an analytical enumeration of all relationships comprising a problem domain – clearly intractable for many applications.

4. The Costs and Benefits of Advanced C4I Applications; CADET Design

Software algorithms that analyze C4I sensor data in an automated and intelligent manner can provide many benefits. However, the time, cost, and expertise needed to develop such applications are major obstacles to developing practical applications that offer a very high return on investment. Solutions that are significantly better are not practical if they cost millions of dollars to develop, and solve inexpensive problems that rarely occur.

CADET represents a significantly new and innovative approach to C4I sensor fusion analysis and application development. Traditionally, sensor fusion algorithms are developed for a well-scoped problem, that is, for a narrowly-defined combination of several factors, including: sensor, mission scenario, operating conditions, and processing task (e.g., target detection). The solution often takes advantage of specific characteristics, nuances, and peculiarities of the narrowly-defined problem so that it gives optimal performance.

The result is an algorithm that works well for the specific scenario. Typically, however, neither the approach used to develop the solution, nor the solution itself, transitions easily to other development tasks. Because the approach typically incorporates a high degree of problem domain-specific information, it does not apply to other sensor fusion and data analysis tasks. Therefore, even for similar problems, the proverbial "wheel" must be largely re-invented for every algorithm development effort. CADET overcomes these limitations in several ways, resulting in many benefits:

- A general approach to sensor data fusion algorithm development, transitioning easily to many other C4I applications.
- Exceptionally-automated, performing many of the development tasks traditionally manually intensive.
- Intelligent, incorporating many advanced algorithms, strategies, and methodologies that have resulted from over decades of data modeling research and development by DoD, AbTech, and many other organizations.

With CADET, the engineer can still apply the same degree of fine-tuning for performance optimization as with traditional approaches – though much faster and easier using a wide array of available algorithms and tools already implemented. The need to develop specialized software will be greatly reduced, but when necessary, the software can be easily integrated because of the plug-in open design of CADET. CADET is highly modular, allowing implementation of application-specific data mining capabilities. "Plug-in" capabilities include data pre-processing, feature extraction, and data post-processing algorithms, alternative modeling; customized queries and reports; and new or modified Expert Mining

Strategies. CADET integrates a relational database management system, AbTech's proprietary Statistical Network modeling algorithm, intuitive user interfaces, and many other capabilities.

CADET Expert Mining Strategies are a defined set of experiments that CADET automatically performs. Expert Mining Strategies are more than merely a script file of keyboard inputs and mouse clicks; they include a degree of intelligence that allows automatic decision mkaing during execution of the trategy. As such, they eliminate much of the manual work typically required and provide a higher level of data mining automation. Two Expert Mining Strategies are described below.

5. Example CADET Application – Information Fusion for Automatic Target Recognition

The goal of information fusion is to process and integrate many different types of numeric data and symbolic information to perform specific situational awareness assessment tasks [9]. There are typically many types of information that operational personnel can use to make war-fighting decisions. The key challenge is determining how to combine these widely varying types of data in a real-time and cost-effective manner. The solution is to use complementary technologies to process sensor and non-sensor data, as shown in Figure 6. Here, we show the synergistic combination of sensor image data with situational information using Statistical Networks and Expert Systems.

To demonstrate CADET ability to build target detection models, consider the Tri-Service Laser Radar (TSLR) sensor data shown in Figure 7. The TSLR sensor provides laser radar (range, velocity, intensity) and passive thermal intensity imagery. These five target views are pixel-registered – a specific pixel coordinate in each of the five individual sensor images corresponds to the same three-dimensional, real-world location. To segment a target from its background, CADET uses a pixel-level classification approach [10, 11]. As shown in Figure 8, CADET allows the user to specify Target and Non-Target pixels, from which databases of example pixels are produced.

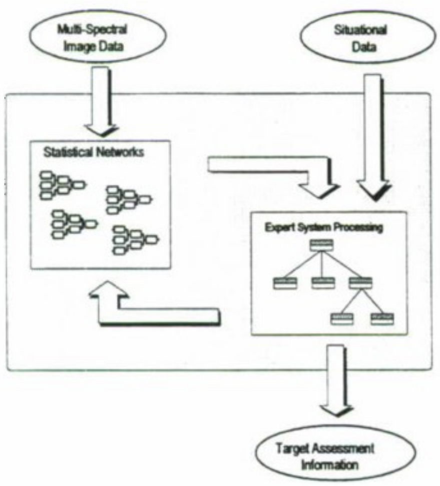


Figure 6 – Integrated Information Processing Approach for Sensor Fusion

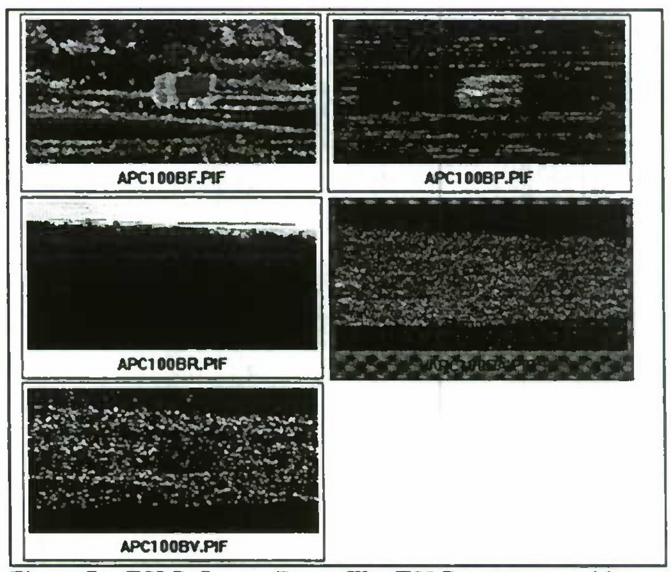


Figure 7 – TSLR Sensor Data. The TSLR sensor provides registered laser radar and thermal target views.

For each pixel, several spatial features are automatically calculated (e.g., the *3x3* window standard deviation). These features provide a characterization of the "behavior" or "texture" of the image immediately about the pixel being classified. The premise is that Target pixels will exhibit characteristics different from Non-Target pixels, which may be subtle. Therefore, pixel characterization features are processed by a CADET Statistical Network which is trained to discriminate between these two pixel types.

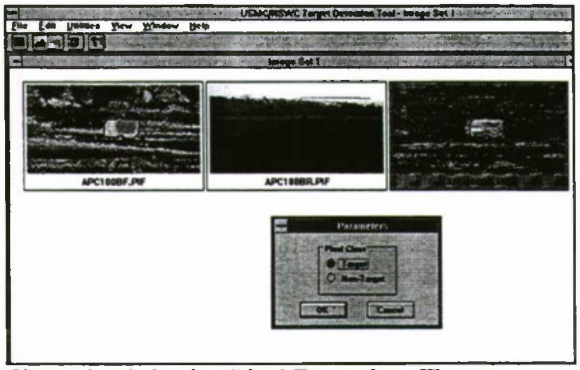


Figure 8 – Selecting Pixel Examples. The user can manually select examples of Target and Non-Target pixels, and define the type of pixel region.

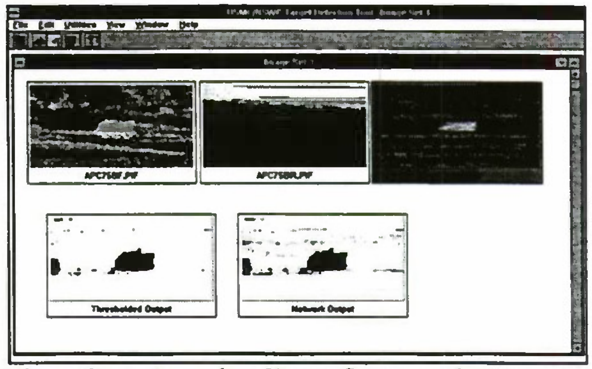


Figure 9 – Independent Target Segmentation Results. The CADET target detection network works very well in segmenting an Image Set *not* used for training.

Figure 9 shows the performance of the network on an independent Image Set, that is, an Image Set that was *not* used during training (and therefore a much more rigorous test). These results show that the detection network is very robust, in that it segments the image extremely well. Note that the small square object at the left side of the image is a target board used for calibration, and as such *should* be detected as a target. Also note, that very little background clutter was detected as target pixels, and that the target itself is well-segmented. These results are even more impressive since the network used here was completely developed in about two hours!

6. Example CADET Application – Belisarius 4 War Game Data Modeling

The Belisarius series was designed to look at the future of the Marine Air Ground Task Force (MAGTF) in view of the Revolution in Military Affairs (RMA) [12]. It was designed to examine how best to cope with, and take advantage of, the RMA while not losing the ability to function in an Operations Other Than War (OOTW) environment. Belisarius 4 deals with an OOTW scenario in the African country of Nigeria, specifically, the political and economical problems which trigger inter-tribal tensions and tensions between the government and the various tribes.

The Belisarius 4 Concepts War Game is played using a complex methodology and a wide range of user resources. These resources include force definitions, weapon system specifications, weather data, and geographic definitions and constraints. Many of the available resources are data tables that define certain capabilities.

For example, a lookup table (LUT) is provided that specifies vertical assault lift capability. It defines the time in minutes to lift a reinforced rifle company (192 Marines and six Dragons) for a given number of aircraft, distance to beach, and aircraft speed. There is a corresponding lookup table for a surface assault lift capability, and one for flying weather visibility percentages. The use of lookup tables is prevalent in this type of war gaming. They facilitate planning during the games moves by providing vital information to ensure that the gaming scenario is realistic.

The Belisarius lookup tables provide "summarized" information to be used for tactical and strategic planning and decision aiding. They are simple in that the indexing variables contain values in very large increments. Lookup table indexing variables are parameters used to define a specific scenario table entry (e.g., time of day, time of year, aircraft speed, number of units). The user "accesses" the LUT with a set of values for each indexing variable, finds the corresponding LUT entry, and uses that output value in the mission planning and decision aiding process.

6.1 LUT Limitations

The primary limitation with LUT is the potential explosion of combinations of input values. A LUT is nothing more than one type of representation of a mathematical function or relationship. For example, consider a function of six input variables, each of which can have one of ten discrete values, as shown in Table 1. A LUT representation of this function requires one million table entries to delineate every combination of inputs. Because many LUTs have many more than merely six input parameters and

more than simply ten possible values, a combinatorial explosion in the number of table entries and the resulting size of the LUT can occur.

Table 1 – Representation of Simple Function with a LUT

Entry #	Param #1	Param #2	Param #3	Param #4	Param #5	Param #6	Table Value
0000001	0.1	0.1	0.1	0.1	0.1	0.1	17.3
0000002	0.1	0.1	0.1	0.1	0.1	0.2	43.7
0000003	0.1	0.1	0.1	0.1	0.1	0.3	2.4
0000004	0.1	0.1	0.1	0.1	0.1	0.4	65.0
0000005	0.1	0.1	0.1	0.1	0.1	0.5	28.4
0000006	0.1	0.1	0.1	0.1	0.1	0.6	10.4
• • •	• • •	• • •	• • •	• • •	• • •	• • •	• • •
999,995	1.0	1.0	1.0	1.0	1.0	0.5	69.9
999,996	1.0	1.0	1.0	1.0	1.0	0.6	44,1
999,997	1.0	1.0	1.0	1.0	1.0	0.7	28.3
999,998	1.0	1.0	1.0	1.0	1.0	0.8	83.5
999,999	1.0	1.0	1.0	1.0	1.0	0.9	77.6
1,000,000	1.0	1.0	1.0	1.0	1.0	1.0	66.3

6.2 Belisarius 4 Flying Weather LUT

For example, consider the Belisarius flying weather visibility percentages lookup tables, provided in Table 2. Its indexing variables include ceiling and visibility, time of day, and time of year. It provides the percentage of time that flying conditions are worse that the specified ceiling and visibility designations. The data was produced by averaging actual data from 16 years of flying observations. For this lookup table, there are 12 time of year increments (the twelve months), eight time of day increments (every three hours), and four ceiling/visibility increments. Therefore, there are only 384 entries in the lookup table.

Table 2 – Belisarius Flying Hours LUT

STA: 653870/DXXX/LOME/TOKOIN, TG-Telecom Summary
LAT: 06 10N LONG: 001 15E ELEV: 82 (ft) 25(m) TYPE: NAVY SMOS V2.0 30091990
20-percent Hours with FLYING WEATHER

CEILING LESS THAN 5000 FT &/OR VISIBILITY LESS THAN 5.00 MILES														
HRS LST	JAN	FEB	MAR	APR	MAY	JUN	JUL	AUG	SEPT	OCT	NOV	DEC	ANN	# YRS
00	42	21	15	10	14	32	41	51	57	36	18	36	31	16
03	53	27	15	13	17	34	57	68	55	25	17	51	36	16
06	88	56	29	19	19	40	63	69	55	43	64	93	54	16
09	74	50	29	24	26	33	54	65	56	41	44	68	47	16
12	41	16	12	8	13	32	42	45	33	16	12	36	26	16
15	30	14	6	5	8	22	22	27	18	9	13	32	18	16
18	39	14	12	8	10	22	23	26	18	16	14	34	20	16
21	27	12	8	10	12	24	30	40	38	20	10	29	22	16
AL L	50	27	16	12	15	30	41	49	41	26	25	48	32	16

CEILING LESS THAN 3000 FT &/OR VISIBILITY LESS THAN 3.00 MILES													
HRS LST	JAN	FEB	MAR	APR	MAY	JUN	JUL	AUG	SEPT	OCT	NOV	DEC	ANN	# YRS
00	28	16	12	8	13	26	31	43	48	29	16	24	25	16
03	36	19	12	11	14	28	44	55	48	19	13	33	28	16
06	75	33	17	13	14	26	42	55	44	23	36	82	39	16
09	54	40	26	21	23	28	42	50	47	34	39	51	38	16
12	26	8	8	6	9	24	22	23	18	12	7	26	16	16
15	18	6	5	4	8	15	11	11	8	7	8	21	10	16
18	24	8	8	7	8	18	12	16	14	13	10	22	13	16
21	16	9	8	9	11	20	22	30	34	18	7	19	17	16
ALL	36	17	12	10	12	23	28	35	32	20	17	36	23	16

CEILING LESS THAN 1000 FT &/OR VISIBILITY LESS THAN 3.00 MILES														
HRS LST	JAN	FEB	MAR	APR	MAY	JUN	JUL	AUG	SEPT	OCT	NOV	DEC	ANN	# YRS
00	19	4	3	0	1	1	1	4	3	2	2	17	5	16
03	23	7	4	1	1	4	2	4	5	1	3	27	7	16
06	70	24	7	4	3	8	15	14	6	6	31	79	23	16
09	32	8	3	*	1	1	2	3	3	1	4	31	8	16
12	23	5	5	*	*	1	1	1	1	1	4	21	5	16
15	18	5	3	1	1	3	3	1	1	1	4	20	5	16
18	22	5	4	1	1	2	2	2	1	2	4	21	6	16
21	11	4	2	1	1	1	1	6	2	2	2	17	4	16
ALL	29	8	4	1	1	2	4	4	3	2	8	30	8	16

CEILING LESS THAN 500 FT &/OR VISIBILITY LESS THAN 1.00 MILES														
HRS LST	JAN	FEB	MAR	APR	MAY	JUN	JUL	AUG	SEPT	OCT	NOV	DEC	ANN	# YRS
00	1	*	*	0	1	*	*	1	*	*	1	2	1	16
03	2	0	0	1	0	3	0	*	1	0	1	4	6	16
06	25	5	1	1	1	1	4	2	1	1	5	25	6	16
09	6	1	1	0	0	0	*	*	*	0	1	6	1	16
12	5	1	1	*	*	*	*	0	*	*	1	5	1	16
15	4	1	*	*	*	1	1	1	1	1	1	4	1	16
18	4	1	1	*	*	*	1	*	*	*	1	5	1	16
21	0	0	*	1	1	1	*	1	*	*	1	3	1	16
ALL	7	1	1	*	*	1	1	1	1	*	1	7	2	16

A potentially more useful lookup table for this application might have weekly increments for time of year (52 potential values), half-hourly increments for time of day (48), and altitude increments of 100 feet (50). Further, consider that the data in Table 2 are for a single geographic location (LAT 06 10N, LONG 001 15E). A more comprehensive model would provide information for multiple locations. The region of interest for Belisarius is an area of dimensions 15° LAT by 20° LONG. Consider the new index variable increments suggested above and 5° LAT-LONG increments. This would require a lookup table with 2,496,000 entries [(52)(48)(50)(4)(5) = 2,496,000], or in other words, a table 6,500 times larger than the sub-tables shown above.

Because this table was built using actual historical data, sufficient data to support this level of accuracy almost certainly does not exist. Also, while this level of accuracy is obviously desirable, it is not practical – it would be extremely tedious to continually search through a LUT of this volume. Therefore, an alternate approach is required. Another challenge is determining table output values for situations where there are no discrete table entries. For example, referring to the tables shown above, for Ceiling Less Than 4000 feet in January at 1030 hours, no actual table entry exists. In some situations, it may suffice to find the closest table entry and use that value. For other applications, however, interpolation between data points is required.

6.3 Modeling Belisarius LUT Data

CADET is ideal for situations where interpolation is required. Here, the reduction of lookup tables to compact algorithmic functions provides a model that can be used more easily over a significantly larger set of scenarios. The resulting model can be used in much the same manner as the original lookup table. However, its advantage is that it can provide outputs for *any* set of input values. For example, a CADET model trained using this data could be used to provide accurate percent flying hours for 2137 hours and a ceiling of 2719 feet altitude. The benefit is more accurate information for war gaming simulation applications.

To build a model of flying weather visibility percentages, we first import the raw LUT data. The CADET interface displaying the imported Belisaurus Flying Time LUT is shown in Figure 10.

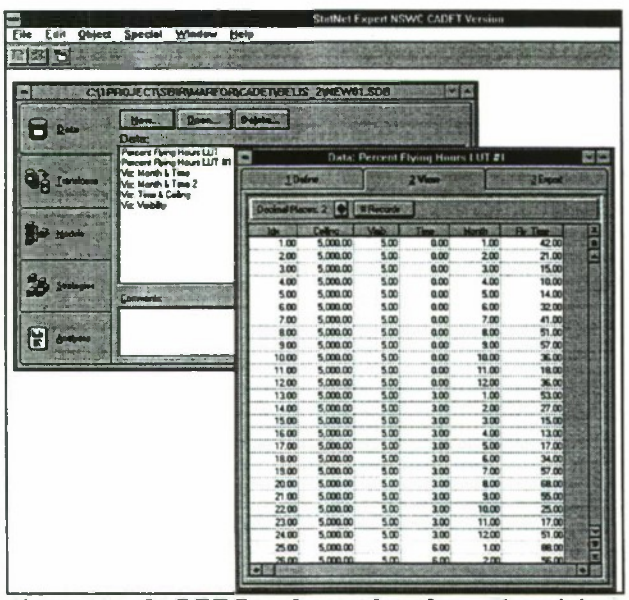


Figure 10 – CADET Data Import Interface. The original Belisarius LUT data.

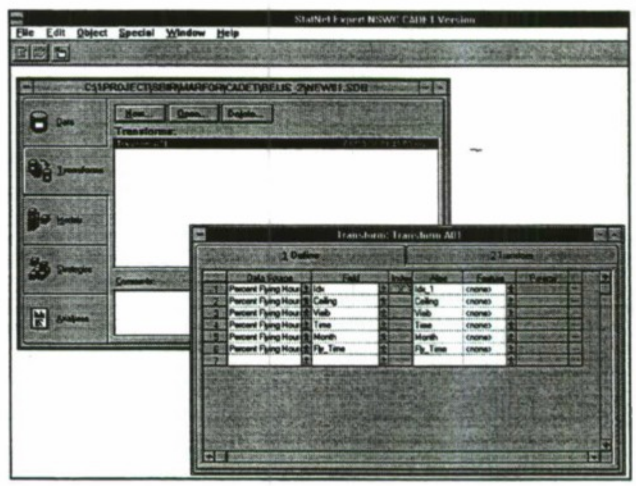


Figure 11 – CADET Transform Interface. CADET allows merging of multiple databases and the definition of feature extraction algorithms to be applied to the original data elements.

We then define a CADET Data Transform, as shown in Figure 11. The purpose of the Data Transform is to extract features from the original variables, merge multiple data tables, and split the data into training and testing subsets. We performed a 75%/25% random data split, resulting in 264 training examples and 88 independent evaluation examples. We then specify a Model Definition by assigning inputs and an output to the data variables. CADET allows the user to choose a data source for a specific model and any combination of inputs and output. Synthesis parameters can be modified to optimize the resulting network model.

CADET then synthesizes a network model using the Statistical Network learning algorithm presented above. Figure 12 shows CADET in the process of synthesizing a network model relating the input parameters Ceiling, Visibility, Time of Day, and Month of Year to the output Percent Flying Time.

The upper right window shows the currently hypothesized network model. CADET will intelligently hypothesize and test hundreds or perhaps thousands of network models during a model synthesis session. The best model found thus far is shown in the upper left window. This network is updated as better models are found by CADET. A set of parameters for each best model is shown in the bottom window. Figure 13 shows the resulting network model, including the equation for the final node.

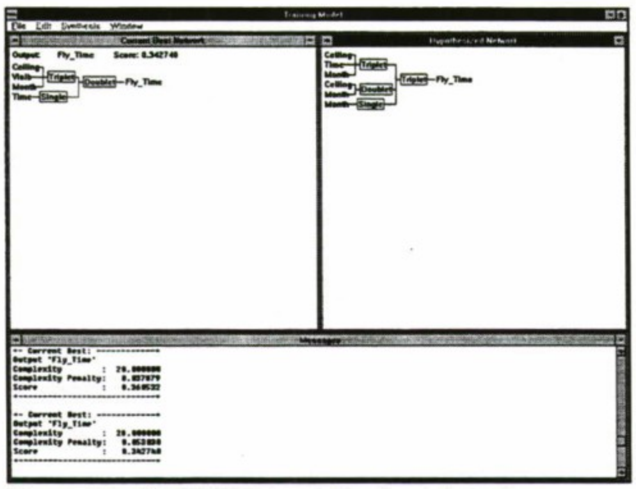


Figure 12 – CADET Network Synthesis Interface. CADET typically hypothesizes and tests thousands of network models during model synthesis. The hypothesized network, current best network, and best network parameters are displayed here.

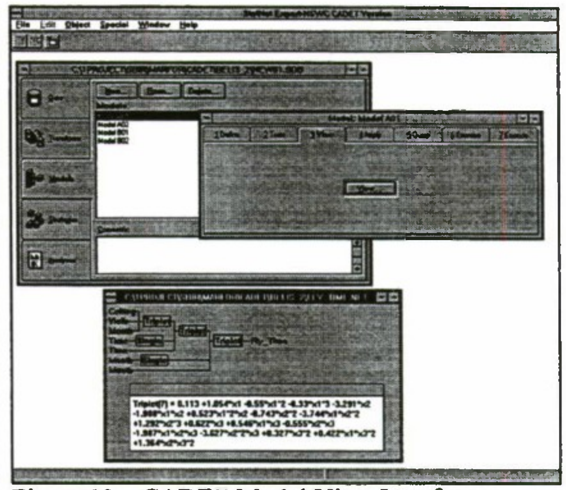


Figure 13 – CADET Model View Interface. CADET allows the user to examine synthesized network models, including the equations for each node.

6.5 Data Enhancement and Feature Extraction

Upon examination of the original CADET data table, it is obvious that a discrete integer representation for the input variables Time and Month (e.g., JAN = 1, FEB = 2, etc.) is not ideal. Time of Day and Month of Year are inherently cyclic parameters. For example, consider the parameter Month. Certainly, January is followed by February, and the value "1" is followed by the value "2" providing an adequate numeric representation for this "symbolic" concept. However, December is followed by January (the end-of-year definition is completely arbitrary, it has no real-world physical significance), but the value "12" (representing December) is not followed by the value "1" (representing January). Assigning sequential integers in this manner "confuses" CADET; it perceives a large discontinuity from December (12) to January (1) where none actually exists.

Table 3 – Modified Belisarius Percent Flying Time Database. A subset of the complete database is shown here.

Ceiling	Visib	Time	Time_0	Time_3	....	Time_18	Time_21	Month	JAN	FEB	....	NOV	DEC	Flying Time
5000	5	0	1	0	....	0	0	1	1	0	....	0	0	42
5000	5	0	1	0	....	0	0	2	0	1	....	0	0	21
5000	5	0	1	0	....	0	0	11	0	0	....	1	0	18
5000	5	0	1	0	....	0	0	12	0	0	....	0	1	36
3000	3	3	0	1	....	0	0	1	1	0	....	0	0	36
3000	3	3	0	1	....	0	0	2	0	1	....	0	0	19
3000	3	3	0	1	....	0	0	11	0	0	....	1	0	13
3000	3	3	0	1	....	0	0	12	0	0	....	0	1	33

Therefore, we represent cyclic parameters with "dummy" or "one-hot" variables as shown in Table 3. Here, twelve new variables are used to represent the actual month for a particular data example. The proper month dummy variable is assigned a value of "1" with the remaining month dummy variables assigned a value of "0." Dummy variables can automatically be created from within CADET.

6.6 Information Content Analysis™ (ICA™) Expert Mining Strategy

The Information Content Analysis algorithm is an Expert Mining Strategy that uses CADET's ability to eliminate inputs, during the network synthesis process, that do not provide useful information to model the output variable. The ICA trains a set of network models at different CPM values to determine which inputs are useful, and to what degree. Variable importance is then quantified. Table 4 shows the results of performing an ICA on the entire data set. It clearly shows that the input variables Month, Ceiling, Time, OCT, and DEC are the largest information contributors. Other variables also contribute information content. Note that more than half of the original variables do not provide any information content.

Table 4 – Results of Information Content Analysis (ICA) Expert Mining Strategy.

Input Variable	Information Content		Input Variable	Information Content		Input Variable	Information Content		Input Variable	Information Content
Month	1.000		Time_9	0.336		Time_3	0.000		MAR	0.000
Ceiling	0.953		Visib	0.065		Time_12	0.000		APR	0.000
Time	0.804		Time_21	0.047		Time_15	0.000		MAY	0.000
OCT	0.804		AUG	0.009		Time_18	0.000		JUN	0.000
DEC	0.804		SEP	0.009		JAN	0.000		JUL	0.000
Time_6	0.336		Time_0	0.000		FEB	0.000		NOV	0.000

6.7 Optimization of Complexity Penalty Multiplier (CPM)

Unlike traditional neural networks, the CADET requires very little "tweaking" with synthesis parameters to optimize performance. However, one CADET synthesis parameter that *is* useful in optimizing network performance is the Complexity Penalty Multiplier [1]. In general, lower CPM values penalize *less* for network complexity, resulting in more complex networks that are very liberal in their use of inputs. Higher CPM values *restrict* network complexity and, therefore, produce less complex networks that are much more discriminating in their use of input variables. We have developed and implemented in CADET an Expert Mining Strategy that finds the optimal value for CPM, using a combination of several statistical optimization techniques.

6.8 Modeling Results

Table 5 defines a set of modeling experiments performed with this CPM Optimization Expert Strategy. The Baseline experiment uses CADET default synthesis parameters, original input variables only, and does not take advantage of Expert Mining Strategies. Experiment 1 uses all input variables (original and dummys), and shows an increase of 9.4% in performance over in terms of Mean Error over the baseline results by optimizing for CPM.

Table 5 – Performance Results for Modeling Experiments

Experiment #	Optimum CPM	Ceiling	Visibility	Time	Time dummys	Month	Month dummys	Mean Error	Maximum Error	R-Squared
Baseline	N/A (CPM = 1.0)	√	√	√		√		5.240	20.442	0.867
1	0.523	√	√	√	√	√	√	4.748	22.438	0.867
2	1.560	√	√		√		√	5.016	27.386	0.821
3	0.434	√	√	√		√		5.287	20.718	0.866
4	0.466	√	√	√			√	5.556	35.517	0.804
5	0.434	√	√		√	√		5.616	27.701	0.799
6	0.487	√	√	√	6, 9, 21	√	AUG SEP OCT DEC	**4.480**	**16.713**	**0.884**

Experiments 2 through 5 use different combinations of the input features, and also optimize for CPM. Experiment 6 takes the best inputs as determined by the ICA analysis discussed above. Here, allowing CADET to focus on fewer inputs often results in better models, as it is not "wasting resources" considering inputs that do not contain useful information. This is true in our case, as the performance of the model produced during Experiment 6 outperforms all other models.

Finally, once the performance of a network has been determined to meet requirements, its network equations and all data preprocessing and feature extraction can automatically be encoded into a stand-alone, embeddable module, which can easily be integrated with existing software applications.

Here, we have taken a relatively sparse LUT relating flying conditions and date/time information to Percent Flying Time, and modeled it to provide a much more robust estimation capability for mission planning and tactical and strategic decision aiding. The resulting model can be used to provide information for a much wider range of scenarios than is provided by the original LUT.

We can use the resulting CADET model to estimate the Percent Flying Time for 4,250 feet altitude, a visibility of 4.5 miles, at 1045 hours on 15 August. The resulting estimate is 47.6%. This scenario could not be computed from the original LUT, as the specified conditions (i.e., index variable values) are not provided. Interpolation between the indexing variables values that do exist is dangerous, as there are very large gaps between these values.

7. Summary and Conclusions

CADET provides a cost-effective data mining capability for a wide range of data modeling opportunities. It is expected that it will play a vital role in support of Current NSWCDD Expeditionary Warfare Thrusts.

References

1. Drake, Keith C., Montgomery, Kenneth M., Kim, Richard Y., and Hanrahan, Patrick E., "Sensor Fusion Through Advanced Computing Techniques," Final Report, contract N00178-95-C-3065, U.S. Naval Surface Warfare Center, Dahlgren Division, Code G32, February 1996.
2. Marine Corps Combat Development Command, FMFRP 14-21,Operational Maneuver from the Sea (coordinating draft), 31 Mar 1995.
3. Holland, T.; Tarr,T.;and Farasie, A., An Artificial Neural System for Target Identification Using IR Imagery Data, Independent Research and Independent Exploratory Development, pp. 5-30 through 5-33, Naval Surface Warfare Center, 1990
4. Marine Corps Research, Development and Acquisition Command, Marine Corps Exploratory Development Program FY 1991 Block Plan, Marine Corps Air-Ground Technology CCIA PE 62131M, 30 Aug1990
5. Holland, O.T.; Poston, W.L.; and Nichols, K.R., "Enhanced Harmogram Analysis Techniques for Extraction of Principal Frequency Components," proceedings of The International Conference on Signal Processing Applications and Technology, Volume 2, pages 906-914 Nov 1992
6. Poston, W.L.; Holland, O.T.; and Nichols K.R., Enhanced Harmogram Analysis Techniques for Extraction of Principal Frequency Components, 1993 NSWCDD Technical Report NSWCDD/TR-92/313,1991
7. Solka, J.L.; Priebe, C.E.; Hayes, I.M.; and Rogers, G.W., Wavelet Transformations fro Helicopter Identification Via Acoustic Signatures, 1994 NSWCDD/TR-93/169,1993.
8. Farlow, S. J.(editor), Self-Organizing Methods in Modeling: GMDH Type Algorithms, Marcel-Dekker, Inc., New York, 1984.
9. Drake, Keith C., and Richard Y. Kim, "Hierarchical Integration of Sensor Data and Contextual Information for Automatic Target Recognition," Applied Intelligence, 5, 269-290, 1995.
10. Drake, Keith C., and Richard Y. Kim, "Pixel-Level Object Segmentation From Multi-Spectral Sensor Imagery, " Proceedings of the SPIE 22nd Applied Imagery Pattern Recognition Workshop, October 1993, Washington, DC.
11. Drake, Keith C., and Richard Y. Kim, "Pixel-Level Object Segmentation Using Moment Invariants As Characterization Features," Proceedings of the SPIE Sensor Fusion and Aerospace Applications II Conference, Orlando, Florida, April 1994.
12. "The Belisarius 4 War Game – Operations Other Than War," Briefing, NSWC.

This page intentionally blank.

AIRCREW AWARENESS OF AIRCRAFT CONDITION

Floyd A. Glenn
CHI Systems, Inc.
Lower Gwynedd, PA

John E. Deaton
CHI Systems, Inc.
Winter Springs, FL

Abstract

Discussions on the concept of situation awareness have focused primarily on pilot awareness of the broad tactical situation, it is also important to consider aircrew awareness of the condition of their own aircraft in a similar sense. Because of the complexity and interconnectedness of aircraft subsystems, considerable interpretation is required to determine the implications of conventional system warnings and to decide on the appropriate course of action to ameliorate a fault condition. This presentation focuses on the case of mechanical fault conditions in rotorcraft. Until recently, only a few simple sensors (e.g., oil pressure and temperature, oil debris sensor) were provided to signal mechanical faults. However, a variety of new sensors and analyzers have become available recently, under the general rubric of Health and Usage Monitoring Systems (HUMS), which purport to provide more sensitive identifications of faults along with the promise of future extensions to fault diagnosis and prognosis as well. The reported research describes efforts to develop aircrew interfaces for mechanical fault information that will be provided from HUMS sensors and processors. A taxonomy of fault management aiding is presented along with a discussion of empirical data on fault management performance of baseline (i.e., non-HUMS) rotorcraft aircrews and prototype HUMS-based aiding and interface designs.

Background

HUMS systems have been developed and implemented by several vendors over the past decade in order to provide improved diagnostic data for ground-based maintenance operations. (Summary reviews of HUMS technology are available in Stevens, Hall, & Smith, 1996; Parry, 1996, Marsh, 1996). The majority of HUMS installations to date have been made by UK and Norwegian operators who provide helicopter ferrying services to North Sea oil platforms. These installations have been made in order to improve flight safety, and have been quite successful in this regard. The U.S. Navy has recently conducted an Air Vehicle Diagnostic Systems (AVDS) program (Chamberlain, 1994) in order to investigate the potential benefits of HUMS technology to improve safety and maintenance in aging rotorcraft fleets (particularly H-53, H-46, and SH-60) and to facilitate the Navy's transition to a paradigm of "condition-based maintenance" (CBM). CBM is intended to replace the current usage-based maintenance policy whereby major maintenance actions and component replacements are scheduled according to recorded aircraft component usage (typically flight hours, operating hours, takeoffs, landings, etc.) as compared to expected component usage life. Since usage-based maintenance policies are necessarily quite conservative, a conversion to a CBM policy is expected to produce significant savings in maintenance costs along with an improvement in aircraft safety.

Current HUMS systems do an excellent job of monitoring and recording usage data in fine detail so as to enable improvements in usage-based maintenance. However, in the area of health monitoring, current HUMS capabilities are more limited. While many mechanical faults can be reliably identified, many others are quite difficult to diagnose. For health monitoring, interest has focused particularly on the use of vibration sensors (i.e., accelerometers) which are located around key drive-train components. Vibration data is especially important for monitoring gearboxes for which (unlike engines) there is very little other indicator data. A variety of techniques have been developed to analyze and interpret HUMS vibration data. One approach employs higher order (e.g., fourth or sixth moment) feature vectors that are derived from the raw vibration data, and then compared with norms for those feature vectors in order to determine when an anomaly exists. Another recently emerging approach uses neural network algorithms to compare raw or processed vibration data with patterns for known faults, and so achieve diagnoses. Both of these approaches require considerable analytical and subjective judgment on the part of the HUMS expert. In the particularly critical and challenging area of rotorcraft gearboxes, high false alarm rates (in the vicinity of 20% to 40%) are reported for returns of non-faulty gearboxes to the manufacturer.

The Opportunity

Whether or not HUMS technology can support a near-term transition to CBM, it appears to have considerable potential for alerting the aircrew of in-flight developments of mechanical faults. This situation, however, presents a rather unique and challenging opportunity in human-centered design. While the underlying technology is evolving rapidly and high levels of uncertainty are associated with many sensor patterns and fault conditions, there are also immediate opportunities to provide valuable lead-time warnings to aircrews of impending catastrophic mechanical failures. While it may be tempting to wait for the HUMS technology to mature before designing an aircrew interface, the potential near-term safety benefit is more compelling. Of course, a thorough cost-benefit analysis must be conducted and an appropriate aircrew interface design must be developed before any HUMS data is presented to the aircrew. The remainder of this paper is concerned with the aircrew interface design process.

The Design Problem

The aircrew interface design problem consists of determining what information to present and how to present it. We will focus mainly on the former of these issues because it is quite difficult and still not fully resolved, and it is logically prerequisite to the latter information presentation issue for which fairly well-established principles and techniques are available.

In order to determine what information should be presented to the aircrew, it is necessary to establish both what information the aircrew needs and wants and also what information can feasibly be generated to satisfy their needs/desires. Since essentially none of the currently generated HUMS outputs are expected to be suitable for aircrew presentation, this process is expected to require iterative refinement. We start by identifying the general types of information that the aircrew needs, then we try to characterize the kinds of information that we think we can generate in the identified categories, then we assess the utility to the aircrew of those specific kinds of information, etc. We assume that all information that might be presented to the aircrew would have to be generated via some kinds of aiding algorithms which would use the HUMS data as their primary inputs. Such algorithms could vary from extremely simple ones like current chip detector lights which just indicate that some anomaly has been detected, to very sophisticated algorithms which would tell the aircrew precisely what to do to respond optimally to the detected problem.

Information Needs Identification

While we did not initially frame this effort in terms of the concept of situation awareness (SA), the relevance of this concept has gradually become evident. Although aircraft cockpit considerations of SA have typically focused on tactical concerns, especially regarding sense of spatial relationships and dynamics with respect to terrain, ground threats and targets, and other aircraft, other applications of SA have addressed the issues of fault alerts (such as for nuclear power plant operators). The connection became particularly evident as information requirements were developed (as discussed in the following sections) and found to closely parallel the SA information levels taxonomy advanced by Endsley (1995). The one significant departure from Endsley's SA framework that we found appropriate to adopt in the present case has been to include action recommendations to the aircrew as a candidate aiding area; except for that addition our information needs determination can be construed almost completely as an SA aiding effort.

Both analytical and empirical approaches to identification of aircrew information needs are warranted. The analytical approach is needed because it is natural to expect aircrews to be somewhat fixated on types of information with which they are already familiar in current cockpit interfaces, so they may not think to ask for some new types of information that might be very helpful. We should attempt analytically to identify all potentially relevant types of information. But it is also important to determine what information the aircrew would actually use if it were available because it could be detrimental to present information that they won't use, and it is likely to be quite costly to create many of the algorithms for generating the candidate aiding information.

Our effort to develop a taxonomy of relevant aircrew information needs began with an earlier investigation of the rotorcraft warning-caution-advisory (WCA) design problem (Hicinbothom et al., 1995). We extended the earlier taxonomy to reflect HUMS considerations to produce the taxonomy illustrated in Table 1. This taxonomy includes two types of information - primary and qualifying information. We assume that the aiding information is constructed on top of a layer of HUMS data and

information. The aiding information serves to convert the HUMS data outputs into information that addresses the various aircrew information needs. We have envisioned a sort of hierarchy of levels of information progressing from anomaly identification through diagnosis, prognosis, impact assessment, and leading ultimately to action recommendation. Also in the aiding information layer, but orthogonal to the preceding hierarchy are several types of qualifying information such as corroboration, criticality, urgency, and uncertainty, which can apply to all levels in the hierarchy.

Primary Information	Qualifying Information
Anomaly detection and localization	Corroboration
Diagnosis	Criticality
Prognosis	Urgency
Root cause determination	Uncertainty
Impact assessment	Precision
Action determination	Completeness

Table 1. Mechanical Fault Alerting Information Taxonomy

Current HUMS systems provide some data on anomaly identification and diagnosis, but none of this is currently considered suitable to pass directly on to the aircrew. There are ongoing research programs to develop diagnostic and prognostic information from vibration data, but the quality and timing of results in these areas are questionable. Various projections have been made of expected developments over the next five years for the precision and reliability of HUMS diagnostic and prognostic capabilities. While HUMS developers may provide some data suitable for cockpit presentation in the next 5 to 10 years, we must assume that most of the information to be offered to the aircrew will have to be developed through some kind of processing that will operate on top of HUMS. We envision production rule (or expert) systems, possibly with some model-based components, to be the primary mechanisms for information generation in most of the identified categories. Although it is certainly desirable that the production rules and models be based as much as possible on solid analytical and empirical data, it is expected that the initial formulation of many of the rules will have to be derived by knowledge elicitation from subject matter experts (i.e., by subjective estimation and speculation). In the area of action determination, we expect that this approach is ultimately the only option, since a value decision must be made about the tradeoffs between the various mission outcome possibilities. For action recommendations, we are essentially producing a sort of electronic NATOPS which must naturally incorporate all existing NATOPS and be fully compatible with it. But NATOPS currently only addresses conditions and anomalies that can be identified with current sensors. HUMS greatly expands the universe of cases that must be addressed.

It is also important to note that not all information needs are necessarily event triggered. Some information requests and system interactions may occur at the initiative of the aircrew for various reasons such as the following:

- command of recording events -- Current HUMS systems are memory limited and cannot record data from all sensors continuously, especially from vibration sensors which produce particularly large data volumes. Intermittent recordings are currently triggered by key flight events (e.g., take-off and landing) as well as by cyclic schedules, but these systems also allow the aircrew to trigger recordings based on aircrew concerns about system performance.
- command to perform special analyses -- As sophisticated options for processing and viewing data continue to develop, we expect the corresponding decisions to be afforded to the aircrew.
- review of subsystem status data/indicators -- We also expect to provide the aircrew with options to examine both current and historical records of all subsystem status indicators.
- trouble reporting -- Trouble reports could be initiated by the aircrew and recorded so as to facilitate ground crew analysis in conjunction with all corresponding HUMS data.

In order to develop an empirical perspective on aircrew information requirements, a study was conducted with CH-46 aircrews at NAS North Island using a motion-base simulator in order to establish baseline performance of aircrews confronted with a variety of mechanical faults and current cockpit instrumentation. Three ten-minute flight segments were conducted by eight variously experienced aircrews (pilot, copilot, and crew chief) in the context of an embassy personnel extraction scenario with one opportunistic search-and-rescue operation. Performance data were collected via videotape, questionnaire, and extensive videotaped debriefs. Objective and subjective data were analyzed to identify areas of attentional focus, types of communications, any evident problems, and types of desired aiding. The results of this study generally supported the above described taxonomy of information needs. Participants reported dividing their attention roughly equally across the categories of display accuracy, diagnosis, prognosis, impact assessment, and action recommendation. However, in classifying recorded voice communications during periods when the malfunction indications were present, we observed many more comments in the category of action recommendation (28%) than in diagnosis (3%) or prognosis (4%), while a substantial number (7%) were concerned with corroboration of indications. Thus, it is clear that aircrews are concerned with all of the identified categories of information at one time or another. More complete reports of this study are in preparation for presentation at an upcoming conference (Deaton et al., 1997 a & b).

Aiding Concept

Figure 1 illustrates the general concept that we have for aiding the aircrew. A special processor will perform information analysis for the aircrew, taking inputs from available HUMS data as well as other avionics and mission plan data. Information will be provided to the aircrew, via some new Warning-Caution-Advisory system interface, in categories such as condition assessment, corroboration, and action recommendations. We also expect that it may be appropriate to offer somewhat different information to different crewstations, such as to support the distinct roles of the pilot and crew chief in the H-46.

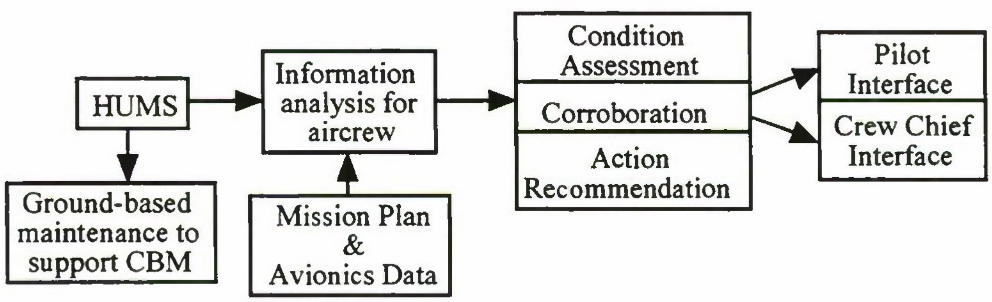


Figure 1. Aircrew Aiding Concept

Conclusions

Continuing efforts are investigating the types of and constraints on information that can feasibly be generated and presented in each of these categories. Decision and information needs are being identified for each of the crew positions in order to develop interface designs that are tailored to the needs of each crewmember. Issues of crew coordination and training for this new alerting and aiding interface are also being addressed. This work is closely coordinated with ONR-sponsored work on diagnostics and prognostics of mechanical systems and associated condition-based maintenance methods.

At the same time there are some substantial challenges. The kind of intelligent WCA aid that we would like to develop can be most easily developed in a new aircraft with advanced computers, reconfigurable displays, and a digital data bus. But the need is greatest in the aging rotorcraft fleets, like the H-46 which doesn't offer the desired infrastructure. Also, the underlying HUMS technology and related technologies for diagnosis and prognosis are developing rapidly. There is a temptation to wait for these technologies to mature, but there is also the concern that the need for aiding is now. Thus, for example, we must weigh the high uncertainties we now have in vibration-based prediction of gearbox

failures against the catastrophic results of in-flight failures. Further, and less problematic, we must recognize that information needs probably vary with many other factors which we have yet to investigate.

Acknowledgments

The work described in this paper has been conducted under Contract Numbers N00014-95-C-0342 and N00014-96-C-0405 with the Office of Naval Research (ONR), with Dr. Harold Hawkins serving as technical monitor. This work has been conducted within a larger ONR sponsored program which is managed by Ms. Jan Dickieson of ONR and supported by Mr. Marty Chamberlain of Jason Associates. This larger program involves the participation of several key people from several different organizations, which include the Naval Air Warfare Center Training Systems Division (Dr. Renee Stout, Mr. Randy Oser, COL Robert Tyler), Pennsylvania State University Applied Research Laboratory (Mr. Bill Nickerson, Mr. Carl Byington), Galaxy Scientific Corporation (Mr. Michael Merriken, Mr. Robert Malone, Mr. Ron Archer), Honeywell Technology Center (Mr. Tom Plocher), and Paul L. Howard Enterprises (Mr. Paul Howard). While all of these organizations and individuals have made important contributions to the overall program, the authors assume full responsibility for the opinions presented in this paper.

References

Chamberlain, M. (1994) U.S. Navy Pursues Air Vehicle Diagnostics Research. Vertiflite, 40 (1), 16-21.

Deaton, J., Glenn, F., Federman, P., Nickerson, G.W., Byington, C., Malone, R., Stout, R., Oser, R., & Tyler, R. (1997a) Mechanical Fault Management in Navy Helicopters. To be presented at the Annual Meeting of the Human Factors and Ergonomics Society, Albuquerque, NM, September 22-26, 1997.

Deaton, J., Glenn, F., Federman, P., Nickerson, G.W., Byington, C., Malone, R., Stout, R., Oser, R., & Tyler, R. (1997b) Aircrew Response Procedures to Inflight Mechanical Emergencies. To be presented at the Annual Meeting of the Human Factors and Ergonomics Society, Albuquerque, NM, September 22-26, 1997.

Endsley, M. (1995) Toward a Theory of Situation Awareness in Dynamic Systems. Human Factors, 37, 32-64

Hicinbothom, J., Weiland, W., Santarelli, T., Fry, C., Voorhees, J., & Zaklad, A. (1995) AMMI: Intelligent Support for Alarm Systems in Advanced Rotorcraft. Presented at the American Helicopter Society 51st Annual Forum, Fort Worth, TX.

Marsh, G. (1996). The future of HUMS. Avionics Magazine, February 1996, 22-27.

Parry, D. (1996). Evaluating HUMS. Avionics Magazine, February 1996, 28-32.

Stevens,, P., Hall, D., & Smith, E. (1996) A Multidisciplinary Research Approach to Rotorcraft Health and Usage Monitoring. Presented at the American Helicopter Society 52nd Annual Forum, Washington, D.C.

This page intentionally blank.

SENSOR-TO-SHOOTER REAL-TIME INFORMATION INTO/OUT OF THE COCKPIT DEMONSTRATIONS

Mark Storer
GDE Systems, Inc.
Mission Planning Systems
San Diego, CA

MAJ Alan Tucker, SAF/SP
MAJ Jeff Johnson, WL/AAZT
Danny Searle, NAWCWPNS
Roger Logan, PMA-233

The United States Air Force and United States Navy are developing Real-Time Information Into/Out-of the Cockpit (RTIC/RTOC) technologies to enhance aircrew situation awareness, increase operational flexibility and increase mission effectiveness against time critical fixed and mobile targets. Real-time information, includes not only essential data for aircrews, but also the weapon systems they carry. Research and development challenges being addressed include: short time critical target timelines, precision responsive flexible targeting, all weather operations, exploitation of real-time information, effective inflight situation awareness, joint service interoperability and cost-effective technology transition. Under an umbrella of Sensor-To-Shooter (STS) activities, GDE Systems, Inc., is supporting the development and integration of onboard/offboard technologies as well as new RTIC/RTOC Concepts of Operations via several multi-phased Research & Development projects, including: SAF/SP Joint Targeting Workstation (JTW), NAWCWPNS Forward Hunter (FH), WL/AAZT Offboard Targeting Experiments (OBTEX), USAF Combat Imagery Support Program (CISP) and NAVAIR Airborne Tactical Information Management System (ATIMS).

Technology development is focused on new offboard and onboard data processing capabilities to rapidly exploit national/tactical imagery and signals intelligence data in near-real-time and coordinate with C^4I operations in order to rapidly prepare and disseminate RTIC/RTOC products. Sample products demonstrated include: strike retasking information, annotated imagery, 2&3-D graphical renderings, mission route updates, precision target aimpoints, guided munition seeker templates, guided munition delivery envelope constraints, threat updates and weather information. During 1995 and 1996, these STS technology efforts were leveraged to demonstrate near-term, mid-term and long-term RTIC/RTOC capabilities in several live exercises, including: Project Strike Phases I & II (SWC/DOZ), Roving Sands '95 (JCS/SPNV), Goldpan 95-1 / 95-2 (ASC/RA), ATIMS Phase II Advanced Information Flow (SPAWAR-32), Army Deeplook '95 & '96 (ARNG), Rapid Retargeting (TACAIR/OTL-166), TMD High Gear (ESC/ZJ), ARPA/DARO Battlefield Awareness Data Dissemination (BADD) and JSF C^4ISR VIP Demonstration

In 1997, new technology development and advanced demonstration preparations are underway. Key technology under development includes Validated Imagery Points (integrated geopositioning tools), multi-source integration (image-to-image source registration and imagery-signals intelligence correlation), Synthetic Aperture Radar Seeker Templates and Strike/Unit Level Softcopy (imagery) manipulation enhancements. Sample demonstrations being supported include Roving Sands '97, USMC Sea Dragon Hunter-Warrior Demonstration, Joint Strike Fighter FTP Demonstration, Goldstrike F-16 RTIC, USMC Silent Fury, ATIMS AV-8B Mission Management Technology Insertion Device Demonstration, USAF Cope Thunder, and F-117A Onboard Mission Management Demonstration.

Operational Assessment projects are also underway to establish RTIC operational requirements, derive an Employment Concept of Operations, facilitate fleet near-term technology transition, and support long-term technology risk reduction. Sample projects include the collaborative ACC-ASC-SAF Joint Endeavor (Bosnia) Rapid Targeting System, NSAWC-NAWCWPNS Rapid Targeting Cell (Fallon NAS) and SAF/SP Mobile National Rapid Targeting Cell (Nellis AFB). These efforts are attempting to orchestrate a joint USAF-USN RTIC test and evaluation methodology as part of routine tactical air combat wing training operations and periodic fleet exercises.

(Reprint of executive summary; formal paper not available)

This page intentionally blank.

ENHANCING SITUATIONAL AWARENESS BY KNOWLEDGE-BASED USER INTERFACES

Dipl.-Ing. G. Dörfel, Dipl.-Ing. H. Distelmaier
Department for Ergonomics and Information Systems (EFS),
Research Institute for Electronics and Mathematics (FFM),
Research Establishment for Applied Natural Sciences (FGAN),
Neuenahrer Str. 20, D-53343 Wachtberg, Germany

INTRODUCTION

Technology pushes for sensor and weapon systems as well as for command, control, and communication (C3) systems have increased the amount and complexity of information available while the time available to process that information has dramatically decreased in modern combat systems. In actual military operations, those responsible for planning and decision making are faced with situations which are characterized by extremely rapid changes in the tactical situation, highly uncertain information, and a large variety of potential situational hypotheses, i.e., in littoral warfare or in crisis and low intensity conflict situations. These decision makers undergo high mental stress due to the need to respond quickly and accurately, or face potentially fatal consequences.

The purpose of the study, as outlined below, is to illustrate that a knowledge-based user assistance with an intelligent and adaptive human-machine interface is a viable approach to overcoming some of these difficulties. Such assistance can support decision makers in performing information processing in all phases of a command and control (C2) cycle, i.e., situation assessment, decision making, and taking action. This approach requires a situation and mission specific allocation of functions between users and machine system components as well as situation and task specific information presentation and decision maker guidance.

The paper starts with a general description of human decision tasks in military situations, and the principles for supporting situational awareness (SA), human decision making and taking action in C3 systems. This is followed by an explanation of a general support concept of a Knowledge-Based User Assistant (KBUA) which consists of a dialogue monitor, a situation monitor, an action planner, and a display manager. Finally, as an extension of the conceptual work, an example is given of prototyping and implementation of the assistant using an air defense task of the Weapon Control Officer (WCO) in the Combat Information Center (CIC) aboard a navy frigate.

SITUATIONAL AWARENESS AND THE DECISION MAKING PROCESS

To cope with complex military situations, Combat Information Center (CIC) personnel must have an overall situational awareness (SA). To establish and maintain SA they must continually acquire all available data and information about the significant factors and conditions including the tactical environment as well as their own combat (sub)system(s). Thus, SA is a prerequisite and forms the basis for decisions at different response levels to achieve goal satisfaction under constraints of situation uncertainty and reaction time availability (Fig. 1). The tasks of collecting and correctly combining different types of data, information, reports, and messages and then drawing accurate conclusions still remain the responsibilities of the CIC personnel who are usually under great strain during tactical operations. Generic cognitive tasks to be performed by means of C3 systems are situation assessment, goal establishment, decision making/action planning, action command, and control of action accomplishment for goal achievement. These cognitive tasks describe the course of activities in military command and control (C2) cycles as well as in nonmilitary decision situations.

Describing the military C2 process from some normative and analytical point of view, Wohl (1981) established a decision taxonomy - the *SHOR* paradigm. His anatomy of the tactical decision process defines four general aspects: *S*timulus, *H*ypothesis, *O*ption, and *R*esponse. The functions required for the stimulus aspect are data gathering, filtering and correlation, as well as aggregation, storage and recall activities. With the results of the stimulus activities, the human operator creates one or more hypotheses referring to the actual situation. Assessing the generated hypotheses, with respect to the data and information available, results in the

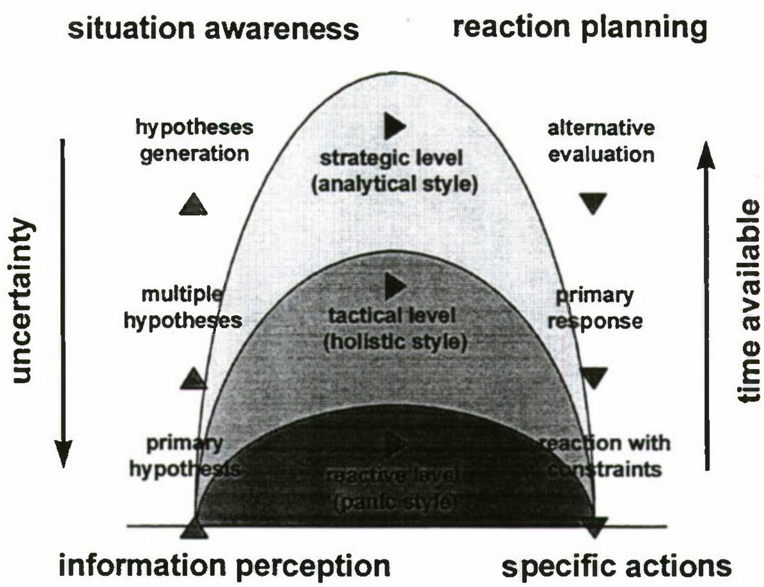


Fig. 1: Decision Making Under Uncertainty and Time Pressure

selection of the hypothesis which yields the strongest confirmation. On the basis of the hypothesis selected, the human operator creates and evaluates option alternatives for goal achievement. The most appropriate option is selected to respond with a specified action on the current situation. Wohl himself stated that some decision situations tend to short-circuit much of the hypothesis and option elements, e.g., in situations with input data of high quality and clearly prescribed options resulting in a classical stimulus-response process.

Studying the decision making process by experts in dynamic and "naturalistic" situations, Klein (1989) found that the majority of time experts focused on situation classification but that a conscious deliberation of solution alternatives was rare. Thus, operators do more than simply collect data about their environment and establish hypotheses on these data or take actions based on only sensor signals. They have to ascertain just the critical features of the situation to understand and classify the situation in the light of their goals, immediately proceeding to action selection.

Endsley (1995) presents SA as the crucial construct on which decision making and performance hinge in complex dynamic situations. The model of situation awareness in dynamic decision making developed by Endsley (1988) emphasizes this aspect. According to Endsley, the first step in achieving SA is to perceive the status, attributes, and dynamics of relevant elements in the environment. Based on these elements, the operator forms a holistic situation comprehension, e.g., a picture of the tactical scenario. Because SA takes into account the dynamics of a situation, it is not only the operators knowledge about the status of the environment at any time, but also includes temporal aspects relating to both the past and the future. Thus, projection of the future actions of the elements, i.e., the future status of the situation, is the third and highly cognitive level of SA. Action selection and performance are separate stages that will proceed from SA.

Besides complexity of dynamic situations, different kinds of information uncertainty and (un)availability will impinge on SA. Thus, the human operator has to cope with, e.g., incomplete, vague, inconsistent or unreliable information and even if the information will be complete, precise, consistent or reliable it might not be available at the correct time or the correct place or in a format suitable to him for perceiving, processing and reasoning. Availability of information is an important aspect of the problem-solving process. In view of the large quantities of information, however, availability alone can no longer be seen as an adequate criterion. Instead, optimum utilization of available information is to be ensured because otherwise the user will be burdened with the difficult task of collecting, storing and cognitively processing relevant information (Woods, 1991).

Wickens (1992) states that the type of information structuring determines how and to what extent information is absorbed and stored. New information interacts with existing mental models of situations and

events. Information representations which are compatible with existing mental models improve information absorption and storage. Thus, a support system should be available for the task of information aggregation (collection, correlation, integration and fusion) as well as for determining type, quantity, form, and organization of the information to be presented.

Results from the literature show differences in the behavior and ability of individuals to perform the different steps of the decision making activities (Hammond, 1988; Crocoll, 1990). Novices will behave more formally and analytically according to the stepwise procedure described by the SHOR model. Experienced human decision makers solve problems primarily by seeing similarities to previously experienced situations. They act more intuitively in a holistic manner modeled by Endsley and Klein. Thus, differences in the cognitive style of the decision making activities will have to influence system design for human interaction. An appropriate support process must be adaptive to the cognitive processes of an individual or a group in terms of the degree to which the system provides the needed information and in a form in which it is compatible to the information-processing abilities of decision makers.

CONCEPT FOR SUPPORTING SA AND HUMAN DECISION MAKING

To support the human operator in SA and decision making within complex systems and situations, adaptive aiding concepts have been developed (Rouse et al., 1988; Rouse, 1991). These concepts to overcoming human deficiencies in information perception and processing have actually been applied as an aid for aircraft pilots as a so-called "pilot's associate" (Rouse et al., 1990; Amalberti and Deblon, 1992; Wittig and Onken, 1992). The basic idea of these concepts is that an overall automation must not be the objective of system development (Bainbridge, 1987). The human operator should be involved in the decision making process as far as his abilities and his performance are sufficient for goal achievement. An aid is provided only to enhance human abilities (e.g., in detecting and evaluating complex patterns or reacting on unforeseen events) and to overcome human deficiencies (e.g., when doing mathematical calculations), i.e., to complement individual human performances. Thus, the human part as an operator and decision maker has to be defined prior to man-machine system design because the overall performance of complex systems strongly depends on human performance, especially when coping with abnormal and unexpected situations.

For the reasons mentioned, the concept of human centered automation recommends a computerized assistant that complements the operator like a human partner. The human user is engaged in a cooperative process in which human and computer assistant both initiate communication, monitor events and perform tasks. The computer assistant does not act as an interface or layer between the user and the application. In fact, the most successful assistant systems are those that do not prohibit the user from taking actions and fulfilling tasks personally, i.e., behaving as a personal assistant that cooperates with the user on the same task. Thus, in parallel to the human operator, the assistant monitors the situation (e.g., states of the system and the environment) and, additionally, operator actions. If the assistant encounters critical situations or inappropriate operator behavior, it may automatically perform some operator-related functions. Faulty behavior of the operator will be classified, announced, and if there is no reaction from the operator, possibly compensated by the assistant. But in any case the user is able to bypass the assistant, so that the responsibility and ultimate decision resides with the human operator.

Together with an interactive graphical user interface the knowledge-based user assistant (KBUA) system forms the knowledge-based user interface (Fig. 2). Information presentation and the user dialog with the C3 system are accomplished via the graphical user interface which acts as the communication tool for the user with assistant and C3 system. The knowledge-based user assistant is not an automation or expert system in the conventional sense but makes the knowledge of domain experts available in the user interface to assist the human operator, e.g., according to situation, mission, task, system states, or operator needs. This aiding or assistance will be attained by a situation- and task-related information presentation (e.g., about the environmental situation or (sub)system states and by warnings or prompts about specific events) as well as by means of operator action guidance according to situation relevant tasks and courses of action.

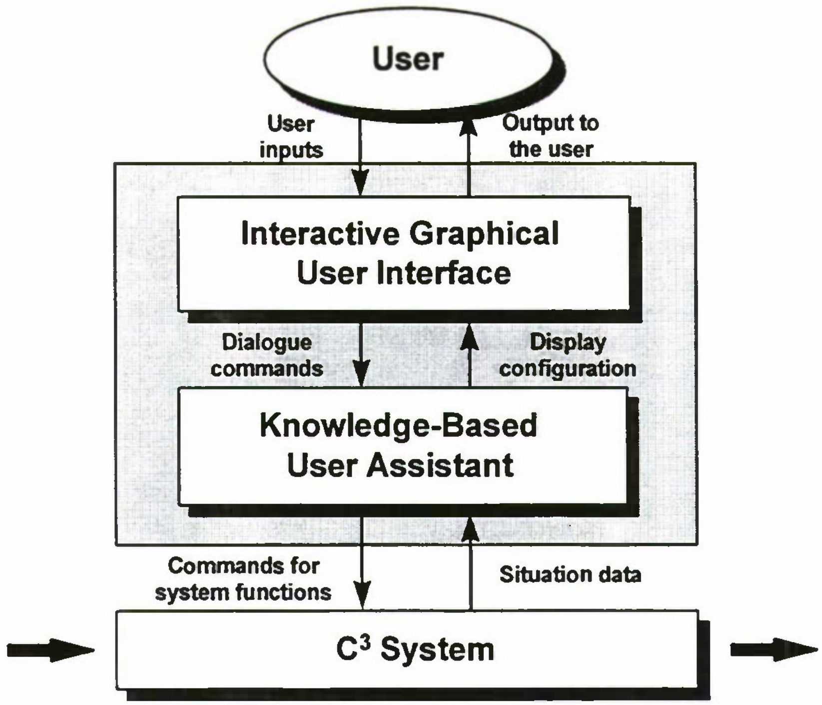


Fig. 2: Structure of a Knowledge-Based User Interface

CONCEPT OF THE KNOWLEDGE-BASED USER ASSISTANT (KBUA)

To support SA and the decision making tasks of the operators in the CIC of a German frigate, a concept for an aiding system has been developed (Berheide et al., 1995). We are proposing the concept of a KBUA system or a knowledge-based user interface, which is a generic concept applicable not only for aiding the members of a naval CIC team in AAW tasks but also for all other missions or the other services, too. The concept itself is like an empty cupboard or shelf where the drawers or boards, i.e., the different functional modules, have to be filled with the special domain knowledge made available by domain experts for special missions and tasks.

The concept of the KBUA system comprises four functional modules representing the situation assessment and the decision making aspects of a C2 process as well as the information presentation and the operator action control aspects. These modules are a Situation Monitor, an Action Planner, a Display Manager, and a Dialogue Monitor as shown in Fig. 3.

The *situation monitor* reviews situation-relevant data and decides about the current state of affairs by combining and evaluating all of the information being available. On the basis of action-relevant situation events detected by predefined filter functions, the *action planner* selects an appropriate course of actions.

Thus, the *action planner* supports the human operator in the decision making process by identifying all actions which are necessary and possible for responding to the currently detected situation event. Dependent on the situation and the time available the action planner decides about the function allocation between human operator and machine system components, i.e., "which" function to be accomplished "by whom" and "when" as well as about the information and action requirements of the human operator, i.e., "what" and "when." In normal, e.g., operational situations without time pressure, all relevant information and all identified and evaluated actions are provided on the user interface by means of the *display manager* and the human operator decides about the actions to be taken. In time critical and complex situations, e.g., multiple air threats in AAW situations, an automatic reaction process may be executed without direct human operator involvement.

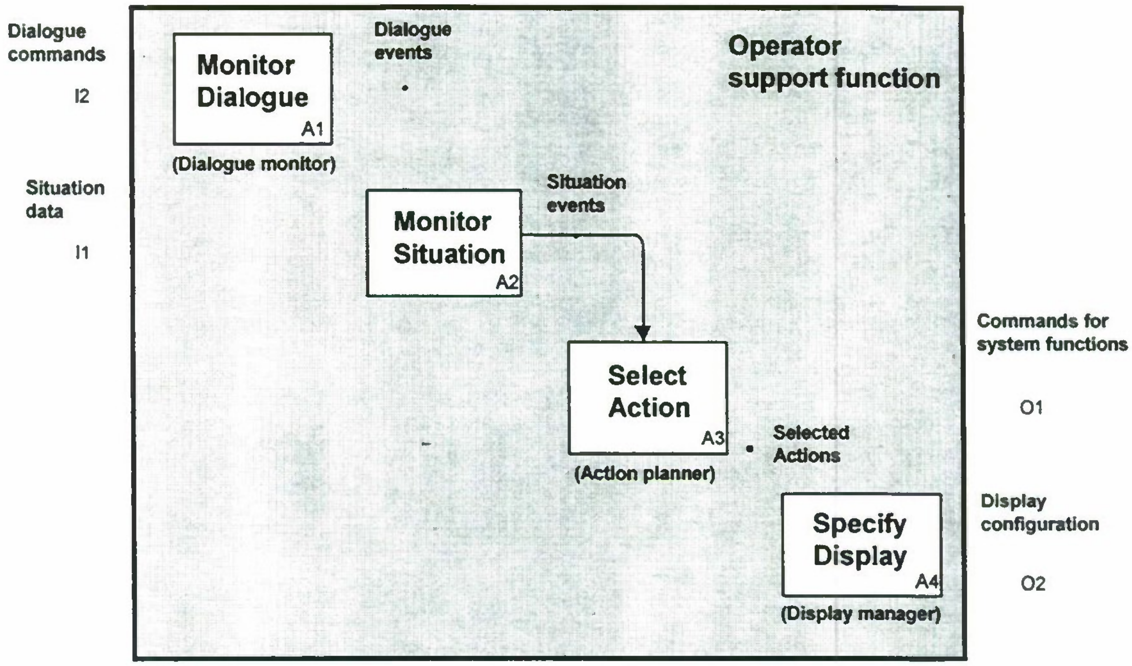


Fig. 3: Functional Components of the Knowledge-Based User Assistant (KBUA)

In this case, the *action planner* generates commands for accomplishing fast automatic system functions. The intent for an automatic reaction is presented to the operator via the *display manager* and may be stopped by VETO or changed by applying own response actions. For each situation and action to be taken, the appropriate display and dialogue elements are stored in the *display manager* as display resources. Information and action possibilities are presented on the graphical user interface by means of the *display manager* deciding "how" to present this necessary information and guidance.
Dialogue events, according to operator command inputs, are provided to the *action planner* by the *dialogue monitor*. The *action planner* compares the actual dialogue commands with those permitted due to the actions to be taken according to the actual situation. Execution of operator dialogue commands not corresponding to the necessary course of actions will be prevented by the *action planner*. Providing a prompt to the human operator supports him in avoiding negative consequences of eventually inappropriate commands and action intents.

DEVELOPING AND IMPLEMENTING A KBUA DEMONSTRATOR

To demonstrate the applicability of the KBUA concept, a knowledge-based user interface has been prototypically developed supporting the Weapon Control Officer (WCO) in the CIC of a navy frigate, especially in threat evaluation and weapon assignment (TEWA) during air defense operations (Berheide et al., 1995). The prototyping approach that we applied started with a relative simple mission and a very simple function model reacting to only a few events. We started our application approach by describing a multi-threat situation in an air defense mission of a ship and identifying relevant mission events and tasks of the WCO. The conditions of relevant events specify the rules of the "Monitor Situation" function of the KBUA. The identified tasks specify the rules of the "Select Action" function of the KBUA and are the basis for designing support functions. To identify functions for supporting the WCO, a hierarchical task analysis has been accomplished, decomposing the demands on the WCO during the described operational situations and resulting in a functional hierarchy with functions of decreasing complexity. As the result of a following cognitive task analysis, information and action requirements have been established for the WCO to accomplish operations on different levels of his air defense task. The information/action requirements identified for each operation are the basis of the "Specify Display" function of the KBUA. These data are used for developing display layouts with the illustration and design tool MACROMIND™. The layouts have been and will be

discussed with experienced users for acceptance and improvements. By decomposing those layouts into their elementary units it was possible to identify required display and dialogue elements as resources for the display manager.

The special scenario for developing a demonstrator for the knowledge-based user interface consisted of an air-defense task of the WCO in a surveillance mission with a multi-threat situation for a ship. In conjunction with defending own ship against an air threat by missiles, the WCO had to prevent non-hostile tracks, e.g., a neutral helicopter from being affected by own weapons (Fig. 4).

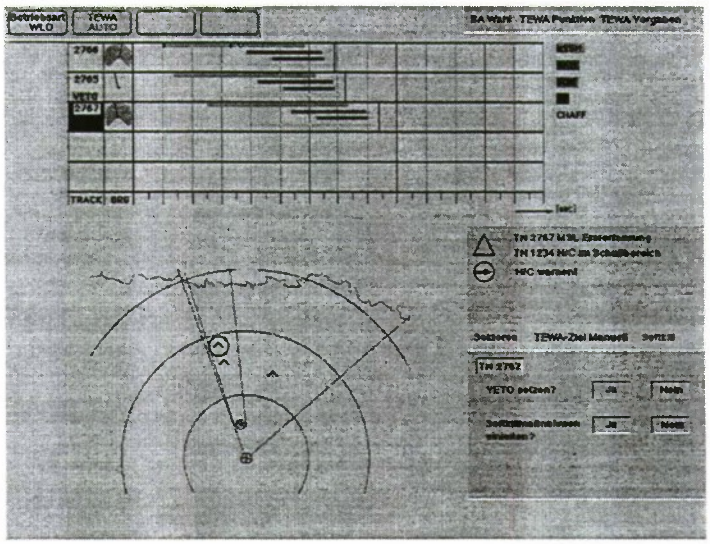


Fig. 4: Design of the Interactive Graphical User Interface for the KBUA Demonstrator

To support the WCO in his decision making activities, he is provided different kinds of event-driven action guidance. When choosing the automatic TEWA process for air-defense he is first requested by the assistant system to establish an "Increased Conflict Area" (ICA). The WCO is pointed to non-hostile air tracks entering this ICA and requested to establish a "Protection Area/Protection Sector" (PA/PS) around those tracks. Whenever a hostile air track, directed to a pre-defined own-ship closest point of approach (CPA) circle and therefore becoming a candidate for automatic weapon assignment, enters the PA/PS of a non-hostile air track the WCO is informed about this event and requested to put a VETO on those hostile air tracks or to use softkill procedures as far as the non-hostile track is in the line of engagement between hostile tracks and own ship. Additionally the WCO is requested to warn the non-hostile air track about his dangerous situation. When the situation changes, because the non-hostile track can evade the dangerous line of engagement by altering its height and/or course, or the hostile track gets much closer to the own ship than the non-hostile track, the WCO is requested do delete the VETO for reactivating the automatic TEWA process against this track. If, because of short ranges and high speeds, it would be time-critical for self defense giving a VETO on the hostile air track, the WCO is informed about this situation and the intended automatic weapon assignment and only requested to warn the non-hostile track.

In addition to the event-driven information, warnings and action guidance possibilities for the WCO, we developed a prototype version of an integrated graphical information display (Distelmaier et al., 1996). Instead of the currently used alphanumeric tables for presenting relevant threat and engagement data, we proposed a dynamic graphical presentation with a time-scale as a basis. This display has moving bars representing time-windows for engagement possibilities for different weapons. Special events like assigned weapon launch, calculated time to intercept, or impact time are indicated by special symbols. An interactive configuration of the graphical display enables the WCO to directly activate special actions on the screen instead of on a separate push-button panel. Fig. 4 shows the prototype design of the display console with the air-defense situation described as a detail of a PPI presentation in the lower left of the display. The lower right of the display shows information in form of textual event presentation in combination with special warning pictographs to enhance SA and presents operator action guidance by preselected action possibilities. By

means of different menu bars the operator is able to bypass the assistant and to respond to the situation by his own intentions. On the top of the display there is the graphical threat and engagement presentation showing actual information for three missile tracks identified by their track numbers. They are ranked by their threat priority. Bearing information of the tracks is visualized as analog presentations. Own ship weapon sectors may be included.

The developed KBUA model is independent of a specific computer language or implementation system. For realizing the knowledge-based user interface demonstrator, we installed the model on a DEC-VAX™ station with the expert system shell SMARTELEMENTS™. Other components of the demonstrator are a pixel-oriented screen with pointing device and a keyboard. The model is implemented with those object-oriented features and rules that SMARTELEMENTS™ offers. The graphical output and dialogue features of SMART-ELEMENTS™ are used as an interactive graphical user interface.

DISCUSSION

Results of decision making and situational awareness research have given rise to the need for decision making aids that can support the human in assessing and reacting to rapidly changing situations. While it may not be currently possible to design a system to address all possible events in highly ambiguous situations, such as those found in littoral operations, it is possible to develop a system to complement the humans ability to perceive novelty and adopt an appropriate response to manage that novelty. The KBUA concept and its implementation demonstrates the potential of such a system.

As pointed out by Zaff, et al. (1993), one of the major shortcomings of many knowledge-based models is the lack of involvement of the users in model development. In the case of the WCO, the user is also the domain expert. When one bases an entire model on the documentation of a domain, (s)he runs the risk of not having the complete scope of the domain nor many characteristics of the user population.

Further development of the KBUA will require additional mission needs analyses starting with a review of C3I operations and systems manuals in preparation for interviewing domain experts. As mentioned above, the experts are the operational WCOs. Interviewing for developing a KBUA needs to be an iterative process with the first interviews being non-directed except for the use of specific operational scenarios , e.g., littoral scenarios can be used to ask what information and data within the scenario does (s)he use for situational assessment and decision making. As we learn more about the WCOs and how they perform their tasks, the more specific the questions can be in successive interviews. After the initial interviews, standardized interviewing techniques will be employed (Sandahl, 1994). The information and data will be analyzed and the results used in further development of the KBUA.

CONCLUSIONS

With the use of the current KBUA demonstrator, the knowledge-based user interface concept for decision making in complex situations was shown to be effective. This was true for SA, as well as for decision making and taking action.

By means of situational user guidance through prompts, warnings, and specific action proposals, operators are alerted to time critical events and the appropriate actions to be taken. This improves SA in addition to decision making and action command. However, the ultimate decision about the action to be taken resides with the operator himself, thus, keeping him as an integral and critical part of the decision making loop.

For specific air-defense tasks, with time and decision-critical response activities, there was a decrease in operator reaction times as well as in operator workload by means of the different support functions.

The integrated and time-based display for threat and engagement information presents the WCO with critically needed information, thus supporting all three levels of situational awareness: perception, comprehension, and projection in the engagement area. The information is task relevant and processed to be compatible with the operator needs, i.e., (s)he can readily assimilate and use the information and data without undue cognitive effort. This improves operator SA and, with the above mentioned situational operator guidance, reduces error in cue awareness and information extraction. Additionally, operator workload is reduced because the information presentation is in terms of the WCO's main goals, i.e., for threat encounters or air defense. Apparently, this is the main reason for the subjective evaluations showing a preference for

time-based, integrated, graphically presented threat and engagement information as opposed to the alphanumeric tables currently used on German Navy ships.

Although the above described approach to developing knowledge-based user interfaces is expensive and time consuming because of the analysis processes, i.e., analyses for system, mission, task, and operator requirements, it yields the advantage of providing a realistic and overall view of the systems operational environment as well as a system design that is closely related to the application and user requirements.

The knowledge-based user interface concept implemented as a specific demonstrator is a very general one and can be applied to different kinds of operator support systems for a variety of different missions and tasks. Human operators will be supported by information presentation and user guidance adapted to mission, situation, task, system status, as well as to their abilities. The object-oriented approach and modular architecture of the concept allows changes and extensions to be made easily to support new and additional functions.

The functionality of the dialogue monitor has not yet been completely realized in the demonstrator implementation. There is the idea to further develop-the dialogue monitor as to dynamically adapt function allocation between human operator and technical system components dependent not only on situational events and task requirements but on operator capabilities and workload, too.

The design of the graphical threat and engagement display is actually used in the operational C2 system software for the German frigates F 123 at the German Naval Command and Control Systems Command's (NavCCSysComd) land-based test site.

The concepts of the KBUA system and the knowledge-based user interface will be further developed. They are actually adopted for developing a support system to aid the CIC team in using Rules of Engagement (ROEs) in crisis and peacekeeping missions.

REFERENCES

Amalberti, R., Deblon, F. (1992). Cognitive Modeling of Fighter Aircraft Process Control: A Step Towards an Intelligent On-board Assistance System. *Int. Journal Man-Machine Studies*, 1992, 36, 639-671.

Bainbridge, L. (1987). Ironies of Automation. In: J. Rasmussen, K. Duncan & J. Leplat (Eds.) *New Technology and Human Error.* New York : John Wiley & Sons Ltd.

Berheide, W., Distelmaier, H., Döring, B. (1995). Adaptive Function Allocation for Situation Assessment and Action Planning in C^3 Systems. Proceedings of the NATO DRG PANEL 8 Workshop on Improving Function Allocation for Integrated Systems Design.

Berheide, W., Boller, H.E., Distelmaier, H., Dörfel, G., Döring, B., Köhler, G., Pontow, T. (1995). *Wissensbasierte Benutzungsschnittstellen für FüWES der Marine.* Forschungsinstitut für Anthropotechnik, Wachtberg-Werthhoven, Germany.

Distelmaier, H., Dörfel, G., Köhler, G., Pontow, T. (1996). *Graphische Darstellung der Bedrohungs- und Bekämpfungsinformation.* Forschungsinstitut für Funk und Mathematik, Wachtberg-Werthhoven, Germany.

Crocoll, W.M. (1990). Status or Recommendations: Selecting the Type of Information for Decision Aiding. University of Massachusetts, AD-A224 133

Endsley, M.R. (1988). Design and Evaluation for Situation Awareness Enhancement. *Proceedings of the Human Factors Society 32nd Annual Meeting*, 1988, pp 97-101.

Endsley, M.R. (1995). Toward a Theory of Situation Awareness in Dynamic Systems. *Human Factors,* 1995, 37(1), 32-64.

Hammond, K.R. (1988). Judgment and Decision making in dynamic tasks. *Information and Decision Technologies*, 14 (1988) 3-14.

Klein, G.A. (1989). Recognition Primed decisions. In W.B. Rouse (Ed.), *Advances in man-machine systems research* (pp. 47-92). Greenwich, CT: JAI

Rouse, W. B., Geddes, N. D., Curry, R. E. (1988). *An Architecture for Intelligent Interfaces: Outline of an Approach to Supporting Operators of Complex Systems.* In: Rouse, W. B. (ed.) Human-Computer Interactions, 1987-1988, Volume 3. Lawrence Erlbaum, Associates, Inc., pp. 87-122.

Rouse, W.B., Geddes, N.D., Hammer, J.M. (1990). Computer-Aided Fighter Pilots. *IEEE*, March 1990, pp. 38-41.

Rouse, W. B. (1991). *Design for Success.* John Wiley & Sons, Inc. New York.

Sandahl, K. (1994). Transforming knowledge from active expert to end-user environment. *Knowledge Acquisition,* Vol. 6, No. 2.

Wickens, C. D. (1992). *Engineering Psychology and Human Performance.* Harper Collins Publishing Inc., New York, NY.

Wittig, T., Onken, R. (1992). Knowledge-Based Cockpit Assistant for Controlled Airspace Flight-Operation. *Preprint 5th IFAC Symposium on Analysis, Design, and Evaluation of Man-Machine Systems*, The Hague, The Netherlands, June 9-11, 1992.

Wohl, J.G. (1981). Force management Decision Requirements for Air Force Tactical Command and Control. *IEEE Transactions on Systems, Man, and Cybernetics*, Vol. SMC-11, No. 9, Sept. 1981, pp 618-639.

Woods, D.D. (1991): The Cognitive Engineering of Problem Representations. In: G.R.S. Weir & J.L. Alty (eds.) *Human-Computer Interaction and Complex Systems.* London: Academic Press.

Zaff, B.S., Mc Neese, M.D., and Snyder, D.E. (1993). Capturing multiple perspectives: a use-centered approach to knowledge and design acquisition. *Knowledge Acquisition,* Vol. 5, No. 1.

This page intentionally blank.

Situational Awareness Training Issues

This page intentionally blank.

UNDERSTANDING AND TRAINING TEAM SITUATIONAL AWARENESS A KNOWLEDGE ELICITATION APPROACH

Renée J. Stout, Ph.D.
Naval Air Warfare Center
Training Systems Division
Orlando, FL

Eduardo Salas, Ph.D.
Naval Air Warfare Center
Training Systems Division
Orlando, FL

Situational Awareness has been identified as a skill that is crucial to the accomplishment of a variety of complex team tasks ranging from military tank platoons tasks, naval command-and-control tasks, fire fighting tasks, and aviation tasks. Its impact within the aviation setting has received a great deal of attention, given that situational awareness is a vital part of every military aviation flight, it plays a key role in aviation safety, mission effectiveness, and operational readiness. Perhaps this is why Endsley (1988b) suggested that situational awareness was the single most important factor for improving mission effectiveness and flight safety for aircrews. For military aviation requiring multi-operator and multi-crew performance, situational awareness is essentially having an understanding of relevant elements in flight and in the mission and predicting their impact on effective mission accomplishment and safe flight. A loss of situational awareness has also been identified as a leading factor in aircraft catastrophes within this military environment. Specifically, an analysis of over 175 Navy/Marine Corps mishaps that were attributed to human error, revealed that the majority were caused by problems in situational awareness (Hartel, Smith, & Prince, 1991).

Despite its recognized importance, situational awareness remains a construct that is poorly understood. A primary reason for this is that the construct has not been adequately defined. One of the most widely cited definitions of situational awareness was provided by Endsley (1987b; 1988b; 1995a), which stated that situational awareness is "the perception of the elements in the environment within a volume of time and space, the comprehension of their meaning, and the projection of their status in the near future" (p. 97). However, this definition has not been met with universal approval. Sarter and Woods (1991) noted that "a commonly accepted definition of situational awareness is still missing" (p. 45). This lack of a clear definition of situational awareness is not trivial in that it hinders the progression of knowledge in this area. This problem is compounded when considering team situational awareness. Specifically, while the criticality of situational awareness in team environments has become obvious, surprisingly few efforts have addressed the team element in situational awareness (Shrestha, Prince, Baker, & Salas, 1995; Salas, Prince, Baker, & Shrestha, 1995). Without a clear understanding of team situational awareness' critical underlying components, tools for assessing team situational awareness and instructional strategies for enhancing team situational awareness cannot be effectively developed. It is critical that training strategies are developed which target required Knowledge, Skills, and Attitudes (KSAs) which comprise team situational awareness. As indicated by Cannon-Bowers, Tannenbaum, Salas, and Volpe (1995) these KSAs will dictate which methods, tools, and strategies are most beneficial for improving team situational awareness.

Several researchers have suggested that the only way to make progress in this area is to start with a systematic, theoretically-based approach at determining the underlying KSAs that contribute to team situational awareness. Recent work by Salas et al. (1995) and Stout, Cannon-Bowers, and Salas (1996) has attempted to provide a conceptual understanding of team situational awareness. Stout et al. presented a theoretical framework of team situational awareness which suggested that team situational awareness is a dynamically changing state, affected by several contextual factors, and developed by individual cue and pattern assessments, team processes, and the shared understanding of the team. They recommended using their framework to guide the investigation of the construct of team situational awareness and the identification of training strategies for enhancing team situational awareness.

Using a theoretical framework as a point of departure, theoretically-grounded measurement tools can be developed and effective predictive validity studies conducted. By showing that there is something in the construct of team situational awareness, which is not available from a direct measure of decision making outcomes or performance, but which then helps to predict decision making and performance, evidence can be revealed that the construct has value. These measures must stem from a comprehensive framework, or else circular arguments of assuming that team situational awareness is being measured,

without theoretical rationale of how it is being measured, will ensue. Once these steps have been accomplished, an understanding may emerge of how team situational awareness impacts training design and delivery. Furthermore, Flach (1995) noted the importance of studying this construct within meaningful, task-rich contexts, which is critical for developing effective training strategies for improving this skill area.

The purpose of this paper is to extend and elaborate upon the work of Stout et al. by describing an approach that is being taken to gain an understanding of the critical components of team situational awareness, using their framework as a point of departure. Specifically, given the strong cognitive influence on team situational awareness noted by Stout et al., this paper will discuss how knowledge elicitation and assessment techniques are being incorporated and interview protocols are being developed, based upon their framework. The utilization of interview protocols that are derived from a theoretical perspective has several advantages and potential payoffs, and these will be discussed. A discussion of the implications for training team situational awareness based upon the approach being taken in this research effort will be provided. The next section provides a brief section of the Stout et al. framework followed by a description of the experimental approach being taken in the current research effort.

A Model of Team Situational Awareness

In attempting to delineate the major components that comprise team situational awareness, Stout et al. (1996) reviewed and integrated the literature in several areas, including research and theory on situational awareness, team training and team performance, and shared mental models, as well as the general literature in cognitive psychology and instructional design. Based upon this review and synthesis, they proposed that shared mental models are a prerequisite for enabling high levels of team situational

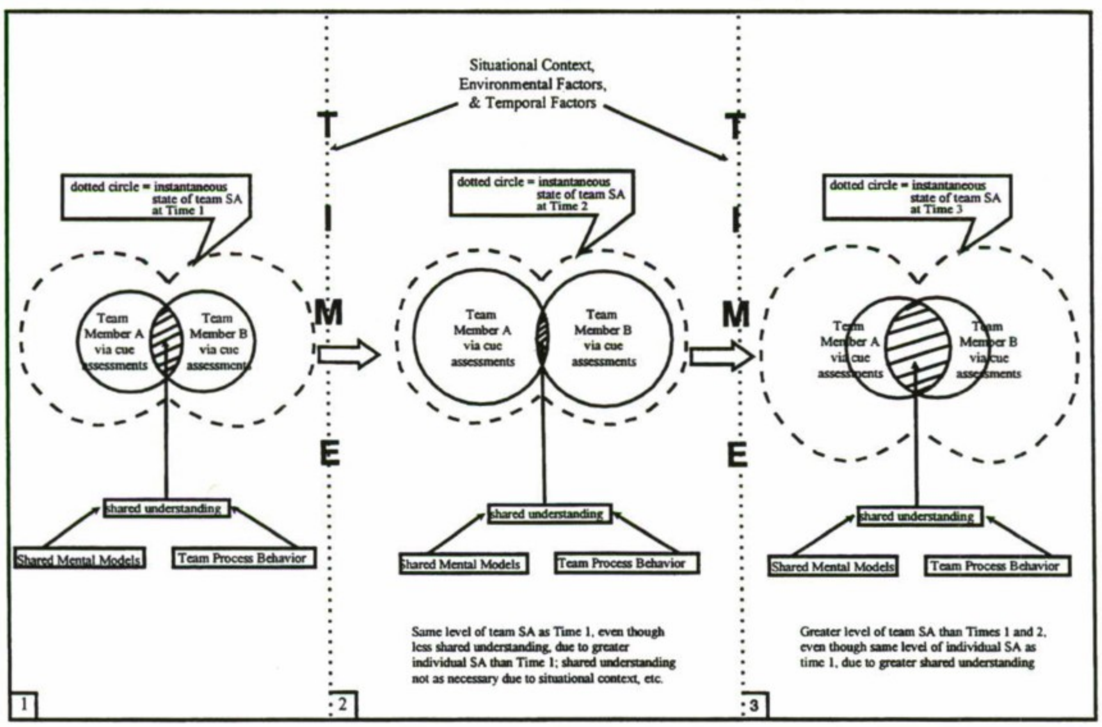


Figure 1. Dynamic team situational awareness

Note. From "The role of shared mental models in developing team situational awareness: Implications for training" by Stout, R. J., Cannon-Bowers, J. A., & Salas, E., 1996. Training Research Journal, 2, p. 104. Reprinted by permission.

awareness in complex, dynamic, and rapidly changing task conditions. They further explained that shared mental models are built through shared declarative knowledge bases (e.g., knowledge about the team task, standard operating procedures, and team member role responsibilities) and procedural knowledge bases

(e.g., knowledge about the steps that are necessary to complete the task and the sequence in which task activities should occur). They reasoned, however, that when the task or mission is high in time pressure, requires a rapid response, and requires team members to adapt to changing task conditions, shared strategic knowledge enables high levels of team situational awareness through utilizing pre-existing declarative and procedural knowledge bases. Strategic mental models contain knowledge of how procedures should be implemented in context, knowledge of what team members should do when events are unexpected or information is absent or ambiguous, and knowledge of actions that should be taken when proposed task solutions fail.

To illustrate their theoretical argument for the importance of shared mental models to the accomplishment and maintenance of team situational awareness, these authors presented a model which explained how shared declarative and procedural knowledge bases are transformed into shared strategic mental models. They then described how shared strategic models are next transformed into team situational awareness, which in turn lead to the taking of appropriate task action. This framework was used to explain task performance across three divergent task conditions, varying in the extent of time pressure and immediacy of required actions when: 1) communications are unrestricted; 2) communications are permissible but must be highly efficient and timely; and 3) communications are virtually precluded.

Also from their review, Stout et al. (1996) proposed an initial model which delineated the major components of team situational awareness, which is shown in Figure 1. Stout et al. suggested that team situational awareness is composed of the culmination of: 1) the situational awareness of each individual team member (based upon the knowledge bases that each team member brings to bear to the task situation and the cue and pattern assessments made by each team member due to pre-existing mental models and bottom-up processing); and 2) the shared understanding of the team (developed through shared mental models and team processes, such as planning, communication, leadership, adaptability/flexibility, and team self correction behaviors). Thus, the three major components of team situational awareness are cue/pattern assessments (via mental models), shared understanding, and team processes. Each of these differentially affect the team situational awareness that is achieved based upon the influence of a host of variables, such as the task characteristics and requirements, the context in which performance occurs, environmental constraints, and temporal factors. Moreover, each component of team situational awareness is differentially required due to the above listed variables. Given this theoretical discussion, we now turn to the approach that we took to accomplish the current research.

Approach

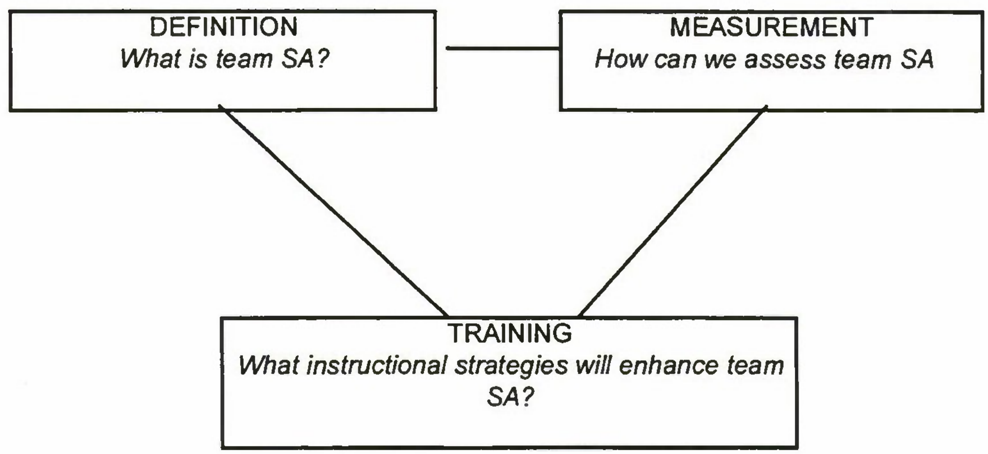


Figure 2. General approach to Team Situational Awareness research

Figure 2 provides a graphical illustration of the approach that is being taken in the current effort. It shows that there are three primary thrusts in our research: 1) definition, or determining what team situational awareness is; 2) measurement, or determining how team situational awareness can be assessed; and 3) training, or determining what instructional strategies can enhance team situational awareness. We argue that these three focal areas are highly interdependent. That is, it is very difficult to specify effective instructional strategies for improving team situational awareness without an adequate understanding of what team situational awareness is and without developing reliable and valid measurement techniques for assessing team situational awareness.

Given an initial understanding of team situational awareness provided by the work of Stout et al. (1996), Figure 3 shows how this framework (Box 1 of Figure 3) can be used to identify measurement criteria (Box 5 of Figure 3) and training strategies for team situational awareness (Box 8 of Figure 3). Box 2 of Figure 3 indicates that we are using the Stout et al. framework as guide for understanding team situational awareness. That is, we are seeking to determine what cues/patterns and team processes are critical for developing and maintaining team situational awareness, and for what elements team members should posses a shared understanding (i.e., the underlying specific KSAs).

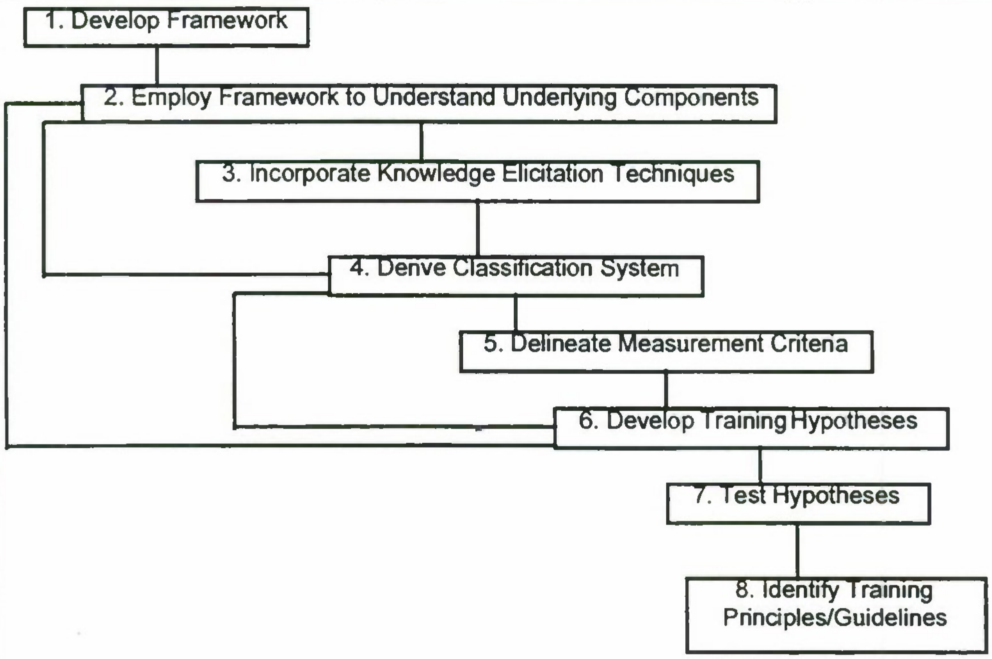


Figure 3. Approach toward incorporating Knowledge Elicitation Techniques to derive measurement criteria and Training strategies for Team Situational Awareness

Given that much of the information that we sought was highly cognitive, we turned to the literature to determine how best to derive this information. Fortunately, the research that has been done in the areas of cognitive engineering and knowledge engineering provided direction in addressing the current research goal. That is, given that much of expert knowledge is automatized, it is extremely difficult to elicit this type of information from experts (Chi, Glaser, & Farr, 1988; Kraiger, Ford, & Salas, 1993) and the use of knowledge elicitation approaches are necessary to "tease out" this knowledge (Cooke, 1994). Thus, the approach taken in this research was to employ the use of a variety of knowledge elicitation techniques to delineate the underlying specific components of team situational awareness (Box 3 in Figure 3). These techniques are derived from knowledge engineering and cognitive engineering perspectives and their advantages and disadvantages have been discussed (see for example, Cooke, 1994). In our research we have employed three primary techniques for uncovering the underlying factors comprising the three major components of team situational awareness: 1) structured interviews with aviators from a variety of naval aviation communities; 2) retrospective interviews of members of an aircrew with use of video of the simulation performance of these aircrews that preceded the interviews; and 3) the scripting of a video of a crew performing both positive and negative examples of team situational awareness as a stimulus material for interviewing aviators. Interview protocols were developed based upon the Stout et al. (1996) framework and tailored to the use of each of these techniques.

We acknowledge that to develop any particular training, one would need to focus on a specific mission and a specific scenario and determine the specific cues/patterns that are most critical for performance, the specific processes that are similarly most effective, and the information that team members should share an understanding of to likewise enable high levels of performance. However, rather than "starting from scratch" for any particular training application, if we can identify a classification system of important cues/patterns, team processes, and elements requiring a shared understanding of the crew, we can use this classification system to guide the development of training in a specific application (Box 4 in Figure 3). That is, if we can identify a set of cue/pattern categories, for example, that are critical across a variety of applications, criteria can be derived for measuring these cues/pattern assessments (Box 5 of Figure 3). Further hypotheses can be derived about how best to train the processing and interpretation of these cues/patterns given underlying cognitive dimensions necessary for utilizing these cue/patterns (Box 6 of Figure 3). These hypotheses can be specifically tested (Box 7 of Figure 3) and then general principles

for training can be generated to drive the development of mission and situation specific training (Box 8 in Figure 3).

In other words, the entire purpose for incorporating knowledge elicitation techniques is to discover the information that should be focused upon in training to therefore make training design and delivery more effective. By uncovering relevant cues/patterns assessments that lead to high levels of team situational awareness and task success, we can design strategies for improving the cue/pattern assessments. Guidelines for scenario development and the use of simulation to optimize the processing of critical cues/patterns can be derived. Decision aiding systems for enhancing team situational awareness can likewise be identified. In addition, the information gained through our approach can lead directly to the specifications for the design of prototypes for training team situational awareness.

Given this discussion as a point of departure, we next turn to describing how interview protocols were derived and linked to the Stout et al. (1996) framework. We will describe this for the structured interviews that we conducted to serve as an example. Following this, we will discuss ways in which we will classify the cues/patterns, team processes, and elements requiring a shared understanding. We will discuss only the classification of cues/patterns as an example.

Identification of cues/patterns, team processes, and shared understanding requirements. Table 1 shows a sample of interview protocols that were developed for use in structured interviews, an indication of whether the protocol was used in general questioning or in conjunction with a paper-based vignette, and the component of team situational awareness (i.e., cue/pattern assessments, team processes, or shared understanding) that each was intended to address.

Table 1
Sample of interview protocols linked to theoretical framework of Stout et al. (1996).

Questions	General	Vignette Specific
Cue/Pattern Assessments		
• What are the most important cues to attend to in any flight?	X	
• What tells you another crew member has lost SA?	X	
• Talk me through the situation and what you are thinking.		X
• What are the three most important cues during this segment?		X
• Describe your expectations for what will happen next.		X
• What cues led to this?		
Team Processes		
• When a fellow crew member has lost SA, how do you help him/her regain it?	X	X
• How can communication among the crew help/hinder team SA?	X	
• What specific aspects of the brief establish good team SA?	X	
• What are the three most important things to communicate to your crew during this segment?		X
• What might a fellow crew member have said that would have been helpful?		X
Shared Understanding		
• Across flights and missions, what do all crewmembers need to be aware of or have a shared understanding of?	X	

• When you have assessed a situation, how do you ensure that other crewmembers have assessed it as you have?	X	
• Are there times when one crew member is doing something and others do not need to know the details? Explain.	X	X
• What might a fellow crew member say or do that would indicate that he/she has not assessed the situation as you have?	X	X

During general questioning, protocols were used, but participants were asked to respond in general across different flights and missions rather than focusing on one particular mission. In other cases, during general questioning, it was necessary for aviators to describe a particular portion of a mission to provide concrete examples during questioning. We also used a paper-based vignette to aid in the interviewing process. In some cases the general questions that we asked were also modified and adapted to the vignette and asked in that context.

Vignette scenario. The vignette scenario described a training mission in which the aviator was told that he/she was the lead aircraft in a flight of two (F-14's for F-14 aviators, for example). The aviator was told that the goal of the training was for him/her to fly a low-level route with predetermined way points to practice low-level navigation. An overview of the mission route, an attached flight plan, a map of the mission area, and a weather brief were all provided. Airspeeds and altitudes were adjusted to make them realistic for low-level navigation for each community interviewed. Times to complete each leg of the route were held constant.

The mission route was divided into five segments, and a description of activities during each leg was described in separate vignette paragraphs. Thus, when starting the route, a description of the first leg was provided and then questions were asked pertaining to that segment. Following this, the paragraph describing the next segment was presented and the questions were repeated for the second segment, and so on. The first segment simply described the crew going feet-dry (i.e., the crew left from the ship and had to enter land) and reaching the point of entry of their flight route. The second segment was also routine but it alerted the crew to potential bad weather conditions by stating that "you notice a squall line approximately 5 miles to the right." During the third segment, the aviator was informed that the crew had arrived at their checkpoint one minute late due to strong head winds, that they had to divert off course due to the weather, and that their next check point was a cluster of three towers on a mountain top. For the fourth segment, the aviator was told that they were still behind time and they were informed of a near miss with a commercial airliner. For the fifth and final segment, the aviator was told that they were still six minutes late, and they were presented with symptoms in their aircraft which indicated potential engine problems.

Aviators were instructed that, while this scenario was paper-based, they were to put themselves into the cockpit as much as possible, see themselves actually flying the mission as described, and respond accordingly. As can be seen in Table 1, questions were intended to probe what the aviator was thinking about, attending to, and talking about. From this, an identification of scenario specific information related to the three major components of team situational awareness could be determined. We turn now to describing our classification of cue/pattern assessments as an example of our approach.

Classification of cue/pattern assessments. To explore how cues/patterns necessary for achieving high levels of team situational awareness can be identified through the knowledge elicitation process, consider the structured interviews that we conducted with aviators from a variety of Naval aviation communities. From these interviews a host of cue/patterns needed in a variety of contexts were verbalized by the aviators that we interviewed. Once we have listed all of the cues/patterns that were discussed by the participants, using a bottom-up approach, we can begin to identify content categories for sorting these cues/patterns. These categories of cues/patterns can add to our definition of team situational awareness by specifying what cues/patterns are important to team situational awareness across a variety of task situations and which are important to particular types of situations. Furthermore, this bottom-up approach may reveal some implications for how we would go about training aviators to utilize these cues/patterns. For example, we may find that some cues/patterns are auditory while others are visual, and based upon

research on training for different information processing modalities, we may differentially train these sets of cues/patterns.

At a deeper level of analysis, we may consider what the source of the cue/pattern is to determine which training strategy to implement. In other words, we might consider who holds the cue and who needs it. Cues and patterns are available from a variety of sources including cockpit instrumentation, the external environment, and fellow crewmembers. In some cases, one particular member of an aircrew may have the only access to the cue/pattern from a nonredundant aircraft system and he/she may be the only one that needs it. In this case, focus might be on individual rather than team training. In another case, one crew member might have access to the cue/pattern but all crewmembers need to be made aware of it. In this case, training might focus on how to efficiently communicate the information to fellow crewmembers. In still other cases, the cue/pattern may be available to all crewmembers, and all crewmembers need it. In these cases, training might focus on helping crewmembers to back each other up when a fellow crew member does not appear to be processing or utilizing the cue effectively.

Finally, at a related yet perhaps deeper level of analysis, according to Stout, Cannon-Bowers, and Salas (in press), it is critical for training to consider what response a set of cues/patterns trigger. Figure 4 illustrates several possible responses that may be necessary as a function of given cues/patterns. These are seen as cognitive processes that update the model of the situation that each team member has and then result in some activity as a result. Cues/patterns are perceived from the situation (e.g., from equipment instrumentation, from the environment, and from fellow team members), and interpreted/projected upon to update one's model of the situation. Based upon situational awareness that is achieved, it may be deemed necessary to (a) do nothing at that point, (b) take a particular action related to individual taskwork (e.g., pull out the emergency procedures checklist), (c) take an action related to teamwork (e.g., copilot back the pilot up on the approach), (d) communicate specific facts (e.g., the copilot tells the pilot that altitude is low), or (e) plan as a team (e.g., the pilot tells the copilot that they will be diverting to an alternate field and says to check close fields for weather and runway lengths). Also, other cognitive activities may occur such as to (f) change one's response set, for example, to use meta-cognitive strategies (e.g., to think to oneself, do I have enough time to plan for what will happen next, or do I need to act); (g) make a decision (e.g., the pilot decides that they need to divert then solicits input regarding his/her decision); or (h) determine that it best is to monitor the situation and look for further cues before acting, communicating, making a decision, or planning. In turn, each activity can update the model of the team member and, indeed, may be required to allow another activity to occur (e.g., if a pilot does not hear from a crew chief that a swimmer is safely aboard, the pilot cannot land the helicopter in the water - a communication or teamwork behavior is required before an individual taskwork action can be made). Also, some activities, such as communicating specific facts and monitoring the situation for changes, influence the perception of particular cues/patterns of cues.

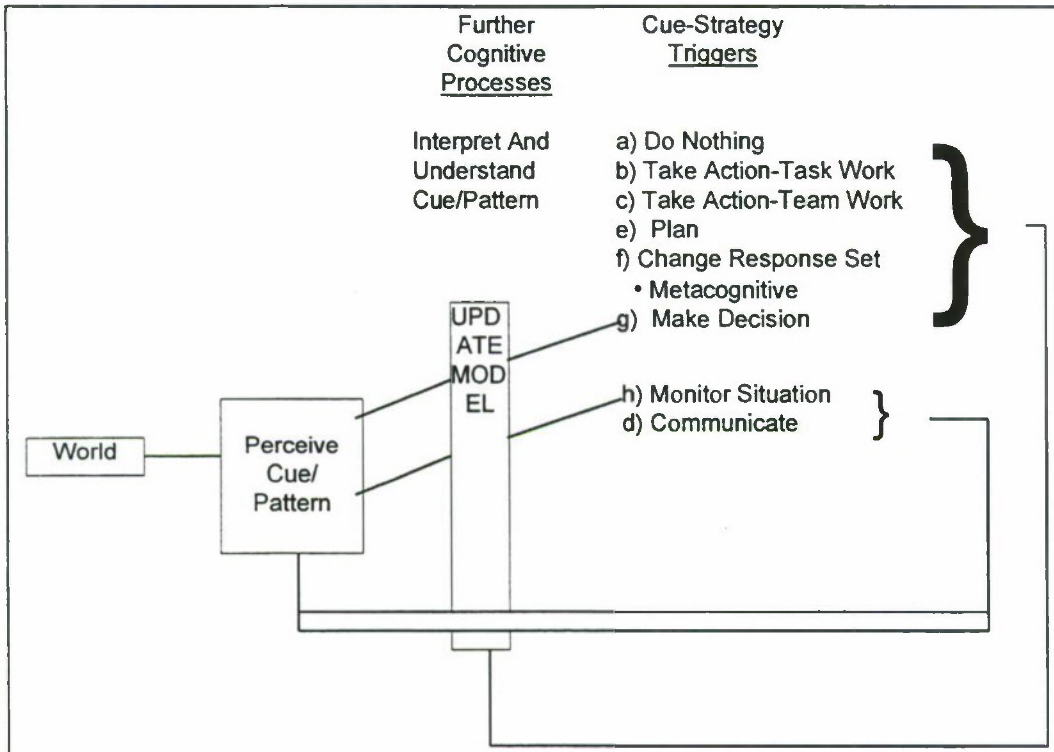


Figure 4. Potential activities resulting from perception and understanding of cue patterns.

Note. From "Team Situational Awareness (SA): Cueing Training" by Stout, R. J., Cannon-Bowers, J. A., & Salas, E., (in preparation). Reprinted by permission.

We argue here that by considering the response that a given cue/pattern activates, we can derive hypotheses for how we might train this association. By combining a bottom-up approach of using our data to delineate the specific cues/patterns important to achieving team situational awareness and the top-down theoretical approach shown in Figure 4, we can identify potentially effective instructional strategies for enhancing team situational awareness.

Thus, through our approach, we will be able to provide guidance for the use of simulation technologies and for scenario design to heighten team situational awareness. By understanding the cues/patterns that are critical to gaining and maintaining team situational awareness, for example, we can begin to specify ways to increase aviators' awareness of these important cues/patterns. Cueing or guiding the trainee's attention to relevant information in their environment can be accomplished in several ways. We refer to this as *cueing training*. For example, cueing can be delivered via system input (e.g., relevant gauges are highlight or faded) or through instructor input (e.g., the instructor describes or points out the relevant information). This can also occur through passive demonstrations or as the team/individual is actively practicing. Finally, this can be accomplished via direct information presentation or via questioning. In each case, the relevant cues come from several sources, including, for example, the cockpit systems (e.g., gauges), the environment, the mission, *and* fellow team members. The actual cueing or prompting to the relevant information can include having the individual/team attend to specific relevant

cues, as well as pointing out effective team processes and task steps to be taken. For team situational awareness, only certain task steps and team processes may be critical to point out.

Four types of cueing training seem particularly relevant to improving team situational awareness (and are summarized in Figure 5):

Passive System Prompting: cueing which occurs through a passive demonstration, such as a videotape or a static system demonstration, which shows the system prompting the relevant information as in active system prompting.

Active System Prompting: cueing which occurs on-line, or as the individual/team is practicing their tasks and which is provided by the system (e.g., via highlighting or fading).

Behavioral Coaching: cueing which occurs through a passive demonstration, such as a videotape or a real time demonstration by an instructor in which the instructor verbalizes the cues he/she *is* attending to in accomplishing the tasks, the relevant processes, and the necessary steps being taken.

Instructor-Guided Practice: cueing which occurs on-line, or as the individual/team is practicing their tasks and which is provided via instructor comments, where the instructor points out the cues to attend to, the processes to attend to, and the necessary steps to take to accomplish the tasks.

In each of these cases, the information can be either directly presented as statements of fact, can be presented by questioning the trainee, or can use a combination of the methods.

Cueing Training

	Passive	**Active**
System	Passive System Prompting	Active System Prompting
Instructor	Behavioral Coaching	Instructor-Guided Practice

Figure 5. Categorization of cueing training methods.

Note. From "Team Situational Awareness (SA): Cueing Training" by Stout, R. J., Cannon-Bowers, J.
A., & Salas, E., (in preparation). Reprinted by permission.

Each of these methods can be used to enhance team situational awareness by pointing out the cues and patterns necessary to make individual situation assessments, and by emphasizing appropriate team processes to utilize within the team setting, given situation assessments. Because each method directs the individual to relevant information in his/her task environment and, thus, helps to build team situational awareness, the choice of one method over another is probably best made on practical grounds considering resources available. However, developmentally, it is argued here, that the strategies of active system prompting and instructor-guided practice are best suited for enhancing strategic knowledge and thereby directly impacting team situational awareness, although the other two strategies provide some aspects of strategic knowledge, yet not as complete due to their passive nature. Each of the strategies can help to develop compatible mental models of the situation at hand. These strategies can also point out relevant information in changing task conditions, to provide contrasts, so that trainees will know what information

to share and will be able to interpret cues in a manner that is both consistent and expected by fellow team members.

With active system prompting and instructor-guided practice, it is hypothesized that team members will form: 1) common explanations of the meaning of task cues; 2) compatible assessments of the situation at hand; 3) common expectations of additional task and information requirements; 4) accurate predictions of team member behavior; and 5) appropriate and expected task strategies. Research is needed to test each of these training strategies to determine their impact on team situational awareness. In addition, research is needed to determine how these strategies should be sequenced in the educational process. That is, behavioral coaching may need to precede instructor-guided practice for optimal impact on team situational awareness. Perhaps most importantly, research is needed to determine what the relevant cues are that should be emphasized during cueing training, which is precisely what our approach will accomplish.

Summary

In this paper we have presented a systematic approach that we are taking to define, measure, and *train* team situational awareness. We have argued that the theoretical framework of Stout et al. (1996) can serve as a guide for further understanding the components of team situational awareness that are in need of training. We have further argued for the importance of incorporating knowledge elicitation techniques for identifying the specific elements (i.e., KSAs) of team situational awareness on which to focus training. Finally, we described how our approach could be used to delineate measurement criteria for team situational awareness and research hypotheses for which instructional strategies might improve team situational awareness. We hope that our work stimulates further research toward training development and design for team situational awareness.

References

Cannon-Bowers, J. A., Tannenbaum, S. I., Salas, E., & Volpe, C. E. (1995). Defining team Competencies and establishing team training requirements. In R. Guzzo & E. Salas (Eds.), Team effectiveness and decision making in organizations (pp. 333-380). San Francisco: Jossey-Bass.

Chi, M., Glaser, R., & Farr, M. (1988). The nature of expertise. Hillsdale, NJ: LEA.

Cooke, N.J. (1994). Varieties of knowledge elicitation techniques. International Journal of Human-Computer Studies, 41, 801-849.

Endsley, M. R. (1987b). The application of human factors to the development of expert systems for advanced cockpits. Proceedings of the Human Factors Society 31st Annual Meeting, 2, 1388-1392.

Endsley, M. R. (1988b). Design and evaluation for situation awareness enhancement. Proceedings of the Human Factors Society 32nd Annual Meeting, 1, 97-101.

Endsley, M. R. (1995a). Toward a theory of situation awareness in dynamic systems. Human Factors, 37 (1), 32-64.

Flach, J. M. (1995). Situation awareness: Proceed with caution. Human Factors, 37 (1), 149-157.

Hartel, C. E., Smith, K., & Prince, C. (1991). Defining aircrew coordination: Searching mishaps for meaning. Paper presented at the Sixth International Symposium on aviation Psychology, Columbus, OH.

Kraiger, K., Ford, J. K., & Salas, E. (1993). Application of cognitive, skill-based, and affective theories of learning outcomes to new methods of training evaluation. Journal of Applied Psychology, 78, 311-328.

Salas, E., Prince, C., Baker, D. P., & Shrestha, L. (1995). Situation awareness in team performance: Implications for measurement and training. Human Factors, 37 (1), 123-136.

Sarter, N. B., & Woods, D. D. (1991). Situation awareness in the advanced commercial aircraft cockpit. Proceedings of the Human Factors Society 34th Annual Meeting, Santa Monica, CA, 21-25.

Shrestha, L. B., Prince, C., Baker, D. P., & Salas, E. (1995). Understanding situation awareness: Concepts, methods, and training. In W. B. Rouse (Ed.), Human/technology interaction in complex systems (Vol. 7; pp. 45-83). Greenwhich, CT: JAI Press.

Stout, R. J., Cannon-Bowers, J. A., & Salas, E. (1996). The role of shared mental models in developing team situational awareness: Implications for training. Training Research Journal, 2, 85-116.

Stout, R. J., Cannon-Bowers, J. A., & Salas, E. (in press). To appear in D. J. Garland and M. R. Endsley (Eds.), Situation awareness analysis and measurement.

Author Notes

The views expressed here are those of the authors and do not reflect the official position of the organization with which the authors are affiliated.

BEYOND PATTERN RECOGNITION: CRITICAL THINKING SKILLS IN SITUATIONAL AWARENESS

Jared T. Freeman, Marvin S. Cohen
& Bryan Thompson
Cognitive Technologies, Inc.
Arlington, Virginia

Daniel Serfaty
Aptima
Burlington, MA

1.1 Introduction

Three cognitive skills play a central role in maintaining situational awareness and tactical decision-making: 1) the ability to quickly recognize important events and recall learned responses to them; 2) the ability to think critically about the meaning of novel events and invent responses to them; and 3) the ability to discern which of these two decision-making strategies is appropriate at a given moment. We describe a model that integrates these processes, training in the latter two skills and the effects of this training on the performance of Navy and Army combat information staff.

1.2 Theoretical foundations

The training we have developed and tested was built upon a model of the role of recognition and metacognition in time-stressed, high risk decision-making environments (Cohen, 1993). This model, in turn, is based on cognitive research indicating the importance of pattern recognition in expert decision-making (Chase & Simon, 1973; Ericsson, et al., 1993; Klein, 1993) and the role of metacognition in monitoring and managing pattern recognition and other skills (Baker, 1995; Cohen, Freeman & Wolf, 1996), as well as theories of the function of mental models in individual and team performance (Rouse, Cannon-Bowers & Salas, 1992; Isaacs & Clark, 1987; Johnson-Laird, 1983).

The following example from the domain of Army battalion operations may help to ground the technical definition of the model. Parts of this example are keyed to the definition, below.

In the heat of battle, a battalion S3 is attempting to locate a friendly recon unit at the request of an Army attack helicopter troop. The helicopters, moving against known enemy troop positions, wish to deconflict the friendly unit from enemy targets. The information streaming to the S3's workstation is voluminous and rich. He receives messages from personnel and systems in the field, the brigade and division above him, the S2 and other members of the staff. From this mass of information he must extract messages from or pertaining to the endangered friendly unit. This is a problem in information filtering, and it is mitigated largely by the clarity of the S3's information retrieval goals (or information goals): he knows what information he needs. As he works, he notices messages from a marginally reliable and poorly positioned scout asserting that enemy wheeled vehicles have just entered the target zone armed with Stinger-like air defense (AD) weapons. Thus, he is also engaged in opportunistic search through the data stream for events that violate his assessment of the situation and his predictions concerning the course of battle (1, below). Realizing that the potential threat posed by the AD weapons is immediate, he quickly relays coordinates of the friendly unit and the enemy AD to the helicopters (in response to their request) and transmits information concerning the AD to friendly artillery units (in anticipation of their requests for help coordinating targeting with the helicopters) (2, below). The S3 also senses that the sightings may have larger tactical implications, and that he has a few moments to investigate those (2, below). In essence, he critiques his assessment of the situation and modifies it to account for the possibility that the AD unit is part of a deliberate defense of a vital enemy point asset, possibly a Command, Control, Communications and Intelligence (C^3I) center concealed near the target zone (3, below). After issuing a call to confirm the sightings, he queries his intelligence assets and his own staff concerning enemy communications, radar emissions and troop movements that might support his suspicion that a C^3I center is near the area. Finally, he advises his own commanding officer to issue a warning order for tank units to prepare to maneuver towards possible vital enemy assets near the observed AD (4, below).

We model this and similar scenarios in the following manner:

1. An officer filters incoming data using either bottom-up, recognition-based faculties or top-down, goal-driven selection criteria that we call information goals. The officer either infers information goals from the interests or responsibilities of others (represented by a mental model of the team), deduces them by reasoning about the tactical situation (represented by a mental model of the situation) or directly

acquires them from commanders' orders and the explicit requests of others.

2. Having selected (filtered) data to which to attend, the officer rapidly evaluates whether there is time and a need to reason deeply about that data. If there is not, the officer executes a well-practiced response and returns his or her attention to the data stream. If there is, the officer proceeds as follows.
3. The officer engages critical thinking skills to interpret the new data and its implications for the situation model. The officer first attempts to formulate arguments with the new data that bear on specific conclusions derived from the situation model. Then the officer critiques the arguments by ferreting out their weaknesses. Three types of weaknesses are hunted. The first is a gap caused by failing to formulate a key argument or a lack of data on which to base an argument. The second is conflict in the conclusions that can be drawn from the available evidence. (E.g., several events may point to one conclusion concerning enemy intent. Other events may point to another conclusion). The third source of weakness is an unreliable argument, which may be based on inaccurate or unreliable data or faulty inference. In sum, the officer uses the data as a lever to pry at weaknesses in the situation model, and the situation model to improve the interpretation of the data. We call this process critical thinking.
4. The officer then acts on the interpreted data by relaying requests, information or recommendations to other officers or by setting new information goals that shape the officer's own information filtering. The better the officer's understanding of the competencies and responsibilities of team members, the better he or she can express and route information, recommendations and requests for information, and the more proactive these actions will be.

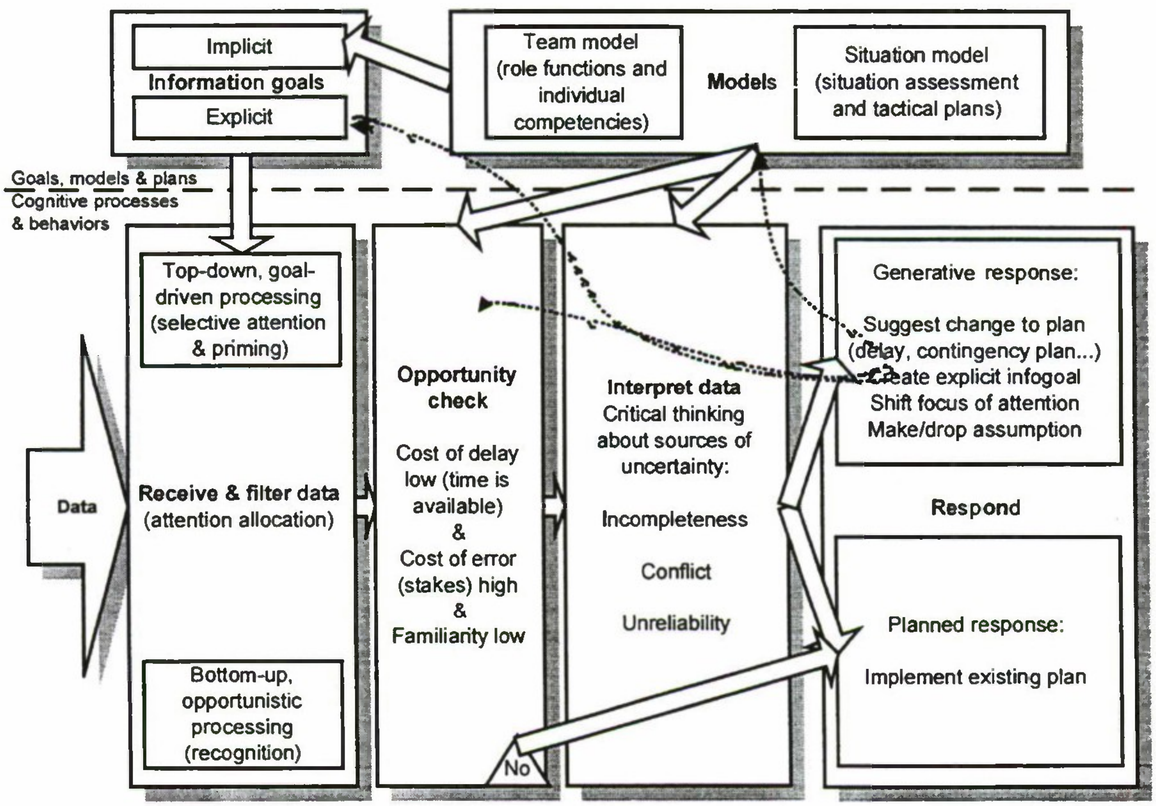


Figure 1: This model of human decision making in an information-processing environment, such as tactical air warfare, integrates recognition and metacognition in a framework that is sensitive to time and risk.

This model has several implications for training officers to maintain situational awareness and make better tactical judgments under stress.

1. Staff who are trained in methods of detecting and handling gaps, conflict and unreliability will identify more, or more crucial weaknesses in the current tactical assessment and plan.
2. Staff who are more sensitive to time constraints, the potential cost of errors and the accuracy with which they recognize a given problem are more likely to correctly decide when to engage critical thinking and when to rely on recognitional responses that are more rapid, but potentially less accurate.
3. The more accurate is a staff officer's mental model of the situation, the better will be the officer's ability to filter useful material from the data stream. If we assume that the staff's commander is the most competent individual to form an accurate assessment, then the commander who communicates his assessment and revisions of it to the staff will indirectly improve their filtering ability.

In short, staff should maintain better situational awareness, form better judgments and take more appropriate actions when they receive instruction in critical thinking and related skills.

1.3 Training

We have developed instruction that addresses each of the training points, above. Here, we describe a version with which train officers who staff the Combat Information Center (CIC) of the Navy's Aegis cruisers. The training is designed for situations where recognition does not function well, that is, where familiar patterns or rules do not fit. For example, some features of the situation may match the standard hostile intent pattern (e.g., an aircraft turning toward own ship from a hostile country). Other features, however, may not match the standard pattern (e.g., its speed is slower than expected), and may even match parts of another pattern (e.g., it is a commercial airliner). In such situations, more-experienced officers explicitly ask themselves how much time they have before they must commit to action. How long can they work the problem until the risk of delay becomes unacceptable? In many situations, when the contact is an immediate threat, critical thinking strategies are not appropriate. The officer may need to immediately take defensive actions. However, it is also important not to undertake an irreversible action, such as engaging a potential threat, before it is necessary to do so. The training focuses on how to use the available time effectively.

The training content is divided into several segments. In each segment, officers listen to a brief verbal presentation of the concepts central to that segment, followed by questions and discussion. They then participate in realistic scenario-based exercises that focus on the trained skill. The training segments concern:

- Creating, testing, and evaluating stories. This section provides an overview of the critical thinking process, called STEP. When experienced decision makers do not recognize situations, they often construct Stories that fit the observed evidence into a robust causal model. A story includes the past, present, and future events that are expected if the best available assessment is true. Decision makers use the story to Test the assessment, by comparing expectations to what is known or observed. When evidence appears to conflict with the assessment, they try to patch up the story by explaining the evidence. They then Evaluate the result; if the patched-up story involves too many unreliable assumptions, they generate alternative assessments and begin the cycle again. In the meantime, they Plan against the possibility that their current best story is wrong by designing contingency plans.
- Hostile-intent stories. Stories contain certain typical components. Knowledge of these components can help decision makers notice and fill gaps in the stories they construct. A particularly important kind of story is built around the assessment of hostile intent. For example, a complete hostile intent story explains why an attack is taking place against a particular target by a particular platform. It also accounts for how that platform has localized the target and is arriving at a position suitable for engaging it. The training teaches officers by practice and example how to reason about such story components and to let the stories guide them to relevant evidence about intent.
- Critiquing stories. After a story is constructed, a decision maker can step back to evaluate its plausibility. This segment of the training introduces a devil's advocate technique for uncovering hidden assumptions in a story and generating alternative interpretations of the evidence. An infallible crystal ball persistently tells the decision maker that the current assessment is wrong, despite the evidence that appears to support it, and asks for an explanation of that evidence. Regardless of how confident decision makers are in their assessments, this technique can successfully alert them to important alternatives. It can also help

them see how conflicting data could fit into a story. In each case, the technique helps decision makers expose and evaluate assumptions underlying their reading of the evidence.

- When to think more. Critical thinking is not always appropriate. Unless three conditions are satisfied, a decision maker should probably act immediately: (1) the risk of delay must be acceptable; (2) the cost of an error if one acts immediately must be high; and (3) the situation must be non-routine or problematic in some way. Training focuses on the way experienced decision makers apply these criteria. For example, they tend to utilize more precise estimates of how much time is available, based on the specifics of the situation. They adopt a longer-term outlook in estimating the costs of an error, and they show greater sensitivity to the mismatch between the situation and familiar patterns.

We have recently added to the training a concept developed in team performance studies for the U.S. Navy (Entin, Serfaty and Deckert, 1994). This is the notion of situation updates, which are periodic statements concerning immediate and potential tactical threats; the updates are made by senior decision-makers to subordinates. Updates concerning immediate threats have a familiar format: the threat is identified and a responsive action is stated. Updates about potential threats concern events that do not readily fit a known pattern. A commander may address such threats with a brief story that accounts for the observations, predicts future events and highlights gaps, conflict and weak assumptions in the story. This addition to the training was tested in research for the Army Research Institute, and is being further tested in team training for Navy Combat Identification Center (CIC) staff.

1.4 Effects of Training

Critical thinking training has now been tested in three studies with 106 military officers or former officers. Studies 1 and 2 were conducted with Naval officers averaging 10 years of military experience. In these studies, participants performed a pretest and posttest on a high-fidelity CIC simulator (DEFTT/TASWIT). Study 2 also employed the simulator to provide practice during training. In both studies, participants responded on paper to questions at three breaks in each test scenario. We asked them to 1) assess the intent of a given track, 2) list other possible intents, 3) defend an assessment other than the one they gave in 1), 4) list the evidence conflicting with a given assessment and defend it, and 5) list actions they would take.

Feature	**Study 1**	**Study 2**	**Study 3**
Location	Surface Warfare Officers School, Newport, RI	Naval Postgraduate School, Monterey, CA	Army Research Institute, Ft. Knox, KY
Participants	60 officers, many with CIC experience	35 officers with highly varied expertise	11 former staff officers
Design	Training (40) vs. control (20) x pretest vs. posttest	Pretest vs posttest	Training (7) vs. control (4)
Duration of experimental session	One day	Five days	One-half day
Duration of training	90 minutes in one day	4 hours over two days	100 minutes in one day
Training tools for executing practice scenarios	Pencil-and-paper	Computer: DEFTT high fidelity CIC simulator	Computer: Training group — Email editor & STIM computerized critical thinking training system Controls — Email editor
Test tools	Computer: DEFTT high fidelity CIC simulator	Computer: DEFTT high fidelity CIC simulator	Computer: Training group — Email editor & STIM Controls — Email editor

Table 1: A comparison of three experimental tests of critical thinking training.

Study 3 was conducted with former Army battalion staff officers. It concerned the impact of critical thinking skills on staff performance under information overload conditions. Participants responded via email to a message stream as a battalion staff operations officer (S3). The stream was periodically paused and participants were asked to make a recommendation concerning a tactical issue, to defend that recommendation and to specify any actions they would take. Controls responded to this break question using the email editor; trained participants responded using a simple drawing tool for representing a decision and arguments in its defense. Responses from both groups were parsed into individual argument points in textual format for analysis. Differences between the designs of the studies are summarized in Table 1.

Training in critical thinking improved situational awareness in all three studies, as indicated by several measures of decision accuracy and the quality of reasoning[1]. In the Navy studies, training improved the accuracy of assessments of intent, relative to those of a subject matter expert, by 79% in study 1 and 35% in study 2 on one of the two test scenarios used in each study. (On the other scenario, no reliable training effect was found). (See Table 2 for statistics.) Contingency planning was 217% more frequent with training on the one study (#1) on which it was measured, suggesting that training helped officers to make or attend to predictions concerning the behavior of suspect tracks, or to find and protect against potential flaws in their assessments. The trained officers cited more evidence in defense of their assessments. This measure improved by 7% in study 1 (not a statistically significant pattern (n.s.), but a strong trend), and 30% in study 2, indicating a more thorough consideration of the events that defined the situation. Trained officers also listed more of the evidence that conflicted with their assessments (52% more in study 1 and 58% more in study 2) and they explained a greater proportion of these conflicts (26% in study 1 (n.s.) vs. 27% in study 2). Thus, training may have helped officers overcome confirmation bias[2] and produce more coherent (less conflicted) arguments. Trained officers also generated more alternative assessments (10% more on study 1 (n.s.) and 41% more on study 2), suggesting that they were better able to avoid tunnel vision with training. Overall, these data indicate that trained officers performed tough critiques of their assessments, yet their confidence in their assessments was 13% higher than that of controls in study 1 (n.s.) and 20% higher after training in study 2. Most officers (73% in study 1 and 71% in study 2) rated the training positively. Ratings were highest among officers with tactical specialties ($F_{1,38} = 4.055, p = 0.051$). Both approval patterns indicate that the training had strong face validity.

In the Army experiment, training in the concepts, above, plus practice with the graphical STIM argument construction tool increased decision accuracy by 34%. The arguments trained participants offered in defense of their decisions were 93% stronger or more persuasive than those of controls, in the estimate of a subject matter expert. This effect was strongest for trained participants with more prior staff training and field experience (*Pearson's* $r = .697, p = 0.082$), suggesting that training in domain-independent critical thinking skills enabled officers to leverage their domain-specific knowledge. The arguments of trained participants were also superior on structural measures, such as counts of supporting evidence, conflicting evidence, gaps, and assumptions. Only trained officers cited conflicting evidence in their arguments and attempted to neutralize it, again indicating that confirmation bias is a significant issue in tactical decision making and that training may help to counter it.

Participants in the Army study also improved on a number of measures of information management. Three related measures of information filtering indicated that controls attended to messages as a function of the rank of the sender, but that trained participants did not do this, presumably focusing on message content, instead ($.05 < p < .09$). With respect to information production, the trained group generated fewer messages ($t_9 = 1.82, p < .05$); these messages were less likely to be simple acts of forwarding ($t_9 = -1.95, p < .05$), more likely to be self-initiated ($t_9 = -1.39, p < .10$) (rather than responses to requests for information) and more likely to convey information rather than to request actions, which might have been performed regardless of the request ($t_9 = -1.95, p < .05$), compared to the messages generated by controls.

---

[1] Differences between groups in the Navy studies were statistically reliable ($p < .05$) unless otherwise stated.

[2] Confirmation bias is the tendency of decision makers to ignore evidence that conflicts with their beliefs and overweight evidence that conflicts with opposing beliefs.

1.5 Conclusion

The research reported above indicates that training in critical thinking skills improves situational awareness and tactical decision-making skills. It presents several interesting avenues for future research and development.

The training should be brought to a wider audience. This can be done in two ways. One is to adapt the demonstration, practice and test material in the training to new domains. It may be possible to complement the generic critical thinking skills with domain-specific instruction concerning, for example, communications protocols and strategies. The second way of reaching new audiences is to automate the training. We have already taken a step in this direction by developing performance measures and measurement instruments for automated training, on behalf of ARI.

Measure	Study 1 (SWOS)	Study 2 (NPS)	Study 3 (ARI)
Accuracy of assessment	79% improvement $\chi^2_2 = 6.337, p = .013$	35% improvement $\chi^2_2 = 6.791, p = .034$	34% improvement $t_9 = -1.467, p = 0.088$
Overall strength of argument (by SME)	[not analyzed]	[not analyzed]	93% improvement $t_9 = -1.647, p = 0.067$
Frequency of contingency planning	217% improvement $F_{1,57} = 8.362; p = .005$	[not analyzed]	[not analyzed]
Number of pieces of evidence raised in defense of an assessment	7% improvement $F_{1,47} = 2.953, p = .092$	30% improvement $t_{33} = 3.807, p = .001$	159% improvement $t_9 = -2.578, p = 0.015$
Number of conflicting pieces of evidence identified	52% improvement $F_{1,55} = 6.236, p = .015$	58% improvement $t_{32} = 5.481, p < .001$	Identified only by trained group
Number of explanations of conflict generated	26% improvement n.s.	27% improvement $t_{32} = 4.920, p < .001$	Generated only by trained group
Number of alternative assessments generated	10% improvement $t_{59} = 1.498, p = .140$	41% improvement $t_{34} = 5.880, p < .001$	300% improvement n.s.
Number of identified gaps in evidence	[not analyzed]	[not analyzed]	642% improvement $t_9 = -2.308, p = 0.023$
Number of assumptions identified	[not analyzed]	[not analyzed]	300% n.s.
Confidence in assessment	13% increase n.s.	20% increase $t_{33} = 1.985, p = .055$	[not analyzed]
Subjective evaluations of training	73% positive	71% positive	[not analyzed]

Table 2: A comparison of results from the three experimental tests of critical thinking training.

Aspects of the training, such as tools for building stories and arguments, might benefit warfighters on the job, particularly in mission planning and possibly in mission execution. To accomplish this, it would be helpful to integrate some of the training concepts with job aids or decision support systems and test them. We have begun to explore this opportunity in a project to design a next-generation CIC training and decision support device.

Finally, while we are pleased to find that critical thinking training benefits individuals, we are intrigued to learn its effects on teams. We have begun to explore this territory in the Army research, involving simulated groups, and with training involving dyads in the Navy CIC setting. However, studies with larger teams might also show benefits of training critical thinking skills.

In sum, training in generic critical thinking skills delivered with domain-specific demonstration, practice and test scenarios has large and reliable effects on situational awareness, and on the accuracy of tactical

reasoning and decision-making.

References

Baker, L. (1985). How do we know when we don't understand? Standards for evaluating text comprehension. In D. L. Forrest-Pressley, G. E. MacKinnon, & T. G. Waller (Eds.), Metacognition, Cognition, and Human Performance (1), 155-205. NY: Academic Press.

Chase, W.G., and Simon, H.A. (1973). The mind's eye in chess. In W.G. Chase (ed.), Visual Information Processing (pp. 215-281). New York: Academic Press.

Cohen, M. S. (1993). The naturalistic basis of decision biases. In G. A. Klein, J. Orasanu, R. Calderwood, & C. E. Zsambok (Eds.), Decision making in action: Models and methods (pp. 51-99). Norwood, NJ: Ablex.

Cohen, M. S., Freeman, J. T., & Wolf, S. (1996). Meta-recognition in time stressed decision making: Recognizing, critiquing, and correcting. Human Factors, 38 (2), pp. 206-219.

Entin, E.E., Serfaty, D., and Deckert, James C. (1994). Team Adaptation and Coordination Training. Technical Report. Burlington, MA: AlphaTech, Inc.

Ericsson, K. Anders, Krampe, Ralf Th., & Tesch-Romer, Clemens. (1993). The role of deliberate practice in the acquisition of expert performance. Psychological Review, 100 (33), 363-406.

Isaacs, E.A., & Clark, H.H. (1987). References in conversation between experts and novices. Journal of Experimental Psychology: General, 116, 26-37.

Johnson-Laird, Philip N. (1983). Mental models. Cambridge, Mass.: Harvard University Press.

Klein, G. A. (1993). A Recognition-Primed Decision (RPD) model of rapid decision making. In G. A. Klein, J. Orasanu, R. Calderwood, & C. E. Zsambok (Eds.), Decision making in action: Models and methods (pp. 138-147). Norwood, NJ: Ablex.

Rouse, William B., Cannon-Bowers, Janis A., and Salas, Eduardo. (1992). The role of mental models in team performance in complex systems. IEEE Transactions on Systems, Man, and Cybernetics (22), 6, 1296-1308.

Acknowledgment

Part of this research was performed under contract to Naval Air Warfare Center / Training Systems Division, in Orlando, FL under various contracts. The opinions expressed in this report are solely those of the authors and do not necessarily reflect the views of the Department of the Navy.

Part of this research was performed under contract to U.S. Army Research Institute for the Behavioral and Social Sciences (ARI) Armored Forces Research Unit at Ft. Knox, KY., contract: #DASW01-97-C-0015. The opinions expressed in this report are solely those of the authors and do not necessarily reflect the views of the Department of the Army.

This page intentionally blank.

WHAT KNOWLEDGE IS SHARED?
TRAINING TRUST IN DISTRIBUTED DECISION MAKING

Marvin S. Cohen
Cognitive Technologies, Inc.
Arlington, VA

Jared T. Freeman
Cognitive Technologies, Inc.
Arlington, VA

Maureen Bergondy
Naval Air Warfare Center
Training Systems Division
Orlando, FL

Successful coordination among team elements may depend on a shared understanding of the situation and team goals. Yet to demand perfectly overlapping knowledge would defeat two essential advantages of teamwork: specialized expertise and distributed workload. For effective training, we need to understand the amount of shared understanding that is appropriate for different kinds of teams and team tasks, along with strategies that compensate for non-shared information.

We are interviewing active-duty flight officers to elicit and model the knowledge used by different air wing elements to plan and execute an air strike, and to identify shared and non-shared knowledge elements. Several different types of knowledge structures have been identified from an analysis of the transcripts of these interviews. Such structures, or mental models, are defined as clusters of topics that recur together in air strike decisions. These mental models include (1) locations and capabilities of enemy forces; (2) friendly plan; (3) friendly higher-level goals, and principles and methods for achieving them; (4) action execution structures, which specify detailed temporal and causal constraints on actions; and (5) evidence interpretation structures, which specify the steps of observation and reasoning from evidence to conclusions.

Different elements of the air wing rely most heavily on different subsets of these knowledge structures. Structures can also differ in their spatial scope (e.g., knowledge of the enemy in only a particular sector versus an overall view of the enemy disposition), in level of detail versus aggregation (e.g., a top-level friendly plan versus detailed plans for its execution), and in their connectedness with other structures (e.g., tracing a friendly plan back to the high-level goals and principles which it reflects, or tracing a picture of the enemy situation back to the observations upon which it is based).

Top-level strike planners and component leaders share knowledge of the friendly strike plan at an aggregated level. However, they differ in their knowledge regarding the overall goals of the campaign (better for top-level planners) and the details of strike execution (better for component leaders). Similarly, strike component leaders and Command and Control (C^2) officers differ in the scope of the air picture they possess (broader for C^2 officers); they also differ in knowledge regarding the way such information is developed (better for C^2 officers). Like component leaders, however, C^2 officers do not have as much knowledge regarding the top-level goals and rationale of the campaign as top-level planners.

Shortfalls in air wing performance could be caused by discrepancies in shared understanding. For example, without an understanding of the overall enemy picture or friendly plan, separate component leaders may develop plans in which they are unable to effectively communicate or coordinate with other friendly units. Without an understanding of top-level goals, component leaders may fail to understand how friendly plans might change due to unexpected events, and C^2 officers may fail to provide the information needed to support such readjustment.

To a degree, these pitfalls of non-shared knowledge are mitigated by strategies in which different units review and evaluate one another's products. For example, top-level leaders evaluate the detailed plans of component teams in terms of higher-level goals. Similarly, C^2 officers critique the detailed plans of component leaders in terms of communication and coordination requirements. New training interventions may facilitate the development of such strategies, to help team elements master the appropriate level of shared knowledge and to monitor for potential problems, without requiring perfectly overlapping information.

(Reprint of executive summary; formal paper not available)

This page intentionally blank.

TACTICAL AVIATION TRAINING: SITUATION AWARENESS AT THE AIR WING LEVEL

Maureen L. Bergondy, Michael Martin
& James Gualtieri
Naval Air Warfare Center
Training Systems Division
Orlando, FL

Jennifer Fowlkes
Summit Technologies
Orlando, FL

To achieve full combat effectiveness, the modern battlefield requires the integration of large tactical aviation teams composed of advanced technology aircraft. The Navy carrier air wing is an example of such a team. Air wings plan, brief, and execute missions that often involve over 100 individuals and more than 30 aircraft that are tasked to complete the subgoals necessary for accomplishment of the mission requirements. Building the situation awareness (SA) of such a large team is challenging for many reasons, not the least of which are issues related to the coordination and integration of the individual team members.

Successful completion of the overall mission requires that a variety of subteams must be orchestrated to fulfill specific mission elements which include strike (STK), command and control (battle space management), the suppression of enemy air defenses (SEAD), combat air patrol, and fighter air interdiction (FTR). During execution, from launch to recovery, these elements must be precisely located in three-dimensional space and their actions carefully timed to achieve the military objective and to protect the aviators. In preparation for these missions, there is a large planning task which must often be performed within a limited time frame.

A small team with representatives from each of the subteams coordinates with a strike leader to develop the concept of operations, an overall mission plan, specific timelines, contingency plans, and detailed sub team plans. This information is then presented in a 50 minute "overall" mission briefing to all members of the strike package. The overall brief is intended to provide the tactical picture, subteam coordination requirements, roles and responsibilities, and the tactical picture of the entire mission execution. The large team then breaks down into its subteams for element briefs of 10-30 minutes. These element briefs are intended to provide detailed subteam coordination and tactical requirements.

To some extent, results within the existing team training and performance domain may be leveraged to support the air wing level research. For example, the concept of mental models has been increasingly invoked as a useful mechanism to explain the performance of teams. In addition, there has been progress in the development of concepts for the training of distributed teams based on a constructivist approach. Both of these areas have produced concepts and guidelines that may be applied to the training of large tactical teams.

However, the nature of large tactical teams suggests that there are many research and training issues that have not been addressed in the current team training and performance literature. For example, with large tactical aviation teams it is probably of greater importance than with other teams to build the big picture or "mission concept" prior to mission execution. Because of the sheer size of the team, and because communications are necessarily restricted during mission execution, it is difficult to regain SA once it is degraded and opportunities for adaptability are limited. Thus, there may be greater emphasis on pre-mission planning and briefing.

Because there is very little opportunity for on-line strategizing during mission execution, implicit coordination is essential. The team member must have a clear understanding of task demands, team coordination requirements, team member responsibilities, contingency plans, and the environmental/mission cues that signal the need for the alternate plan. We contend that the plan and the brief provide the mechanisms to set up effective team coordination by building mental models of the tasks (missions), interdependencies, and contingencies. Effective planning and briefing are expected to improve implicit coordination and thus, mission performance of the large distributed team (Figure 1).

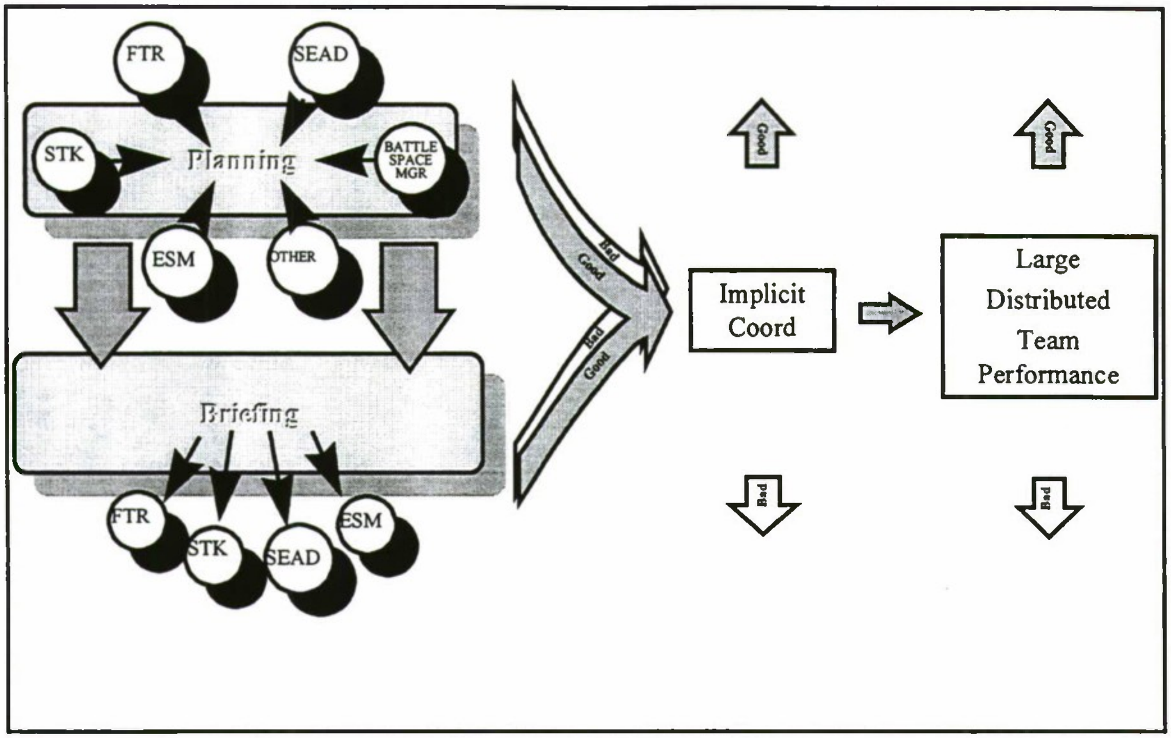


Figure 2. Conceptual model of factors affecting large distributed team performance.

The implicit coordination strategies include all team members monitoring mission progress; recognizing deviations; and correcting team and individual performance through self correcting behaviors, limited communications, or executing pre-planned contingency operations. To execute these strategies in this complex tactical environment, the requirements for individual and team SA are very high.

In this paper we will present some preliminary data that begin to support the relationship between planning, briefing, and execution and raise some questions for future research to address the need for training specific knowledge and skills to improve implicit coordination, situation awareness, and mission performance of large distributed tactical teams.

APPROACH

Two types of data were collected across several air wing training events: direct behavioral observations of live flight training events and training utility questionnaires.

Critical Incident Database

A critical incident data base is developed by observing line flights of five carrier air wing training detachments to an instrumented TACTS range, incidents related to loss of team or individual SA during mission execution (e.g., spillouts, timeline errors, missed tactical calls) as well as, coordination behaviors were identified. The antecedents of these problems/incidents were then traced to planning and briefing issues and observed performance trends across training. Critical incidents were based on direct observation of line flight training events using an instrumented TACTS range. SME Instructors provided identification of effective/ineffective performance from planning, briefing, and mission execution.

<u>Post-detachment Training Utility Questionnaire</u>

Additionally, following the three-week training evolution a training-utility questionnaire was administered to 101 participants across two of the air wings. Participants were representative of air wing composition across, rank, aircraft type, and number of flight hours. This questionnaire

was designed to tap three aspects of the effectiveness of the training detachments mental model development, team processes enhancement, and the impact on OR benefit of specific training phases (Table 1).

Table 1. CATEGORIES OF TRAINING UTILITY QUESTIONARE ITEMS

MENTAL MODELS	TEAM PROCESSES	TRAINING PHASES
Mission Types	Tactical Skills	Pretraining Preparation
Mission Characteristics	Situation Awareness	Planning Missions
Own Aircraft Roles	Threat Awareness	Briefing Missions
Other Platform Capabilities	Integration/Coordination	Execution (Own Proficiency)
Asset Relationships	Flight Discipline	Execution (Air wing Proficiency)
Other Squadron Procedures	Contingency Planning	Mission Debrief
Personnel Familiarity	Decision Making	
	Communications	

RESULTS

Critical Incident Data Base

The critical incident data base has been preliminarily categorized into eight major classes of behaviors based upon subject matter expert inputs: 1) integration, 2) contingency planning, 3) communications, 4) flight discipline, 5) battle space management, 6) target area tactics, 7) suppression of enemy air defenses, and 8) flight tactics. Figure 2 shows the relative proportion of errors by category.

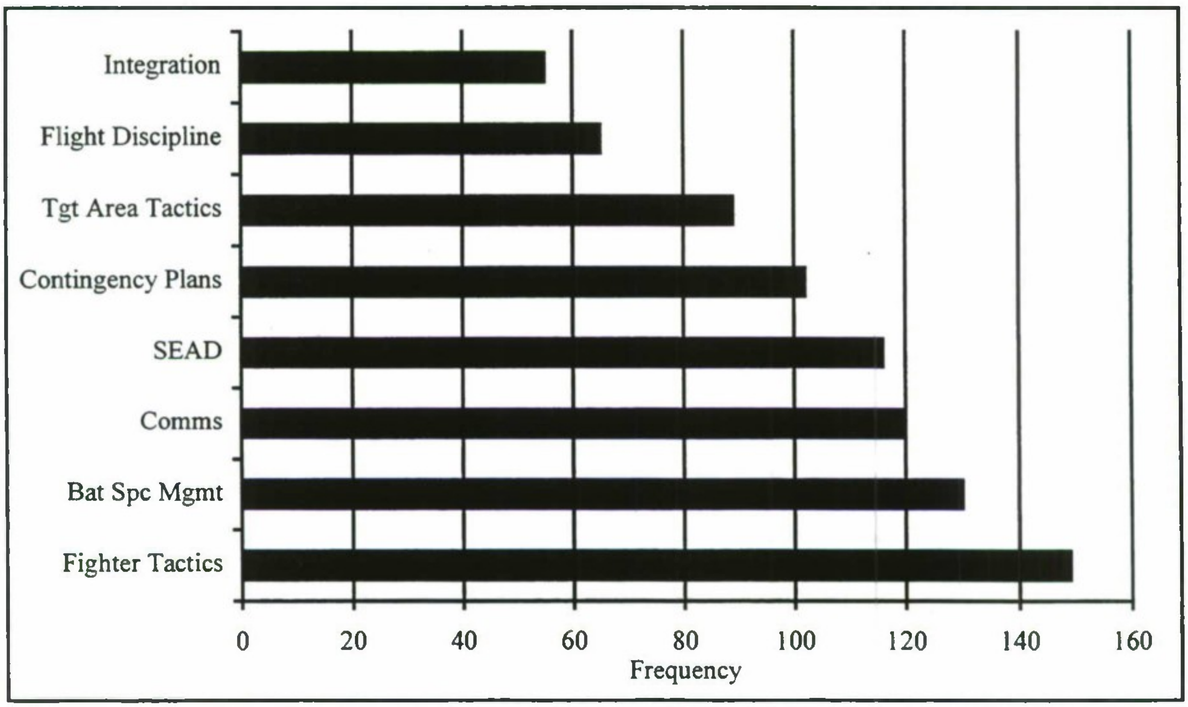


Figure 2. Relative proportion of errors by category.

We then examined the phase of the mission in which each of the critical incidents occurred to begin looking at the relationships of the plan, brief, and execution. Communication is extremely important for building and maintaining SA. Figure 3 shows the frequency of effective communication behaviors and errors across the different phases of the mission.

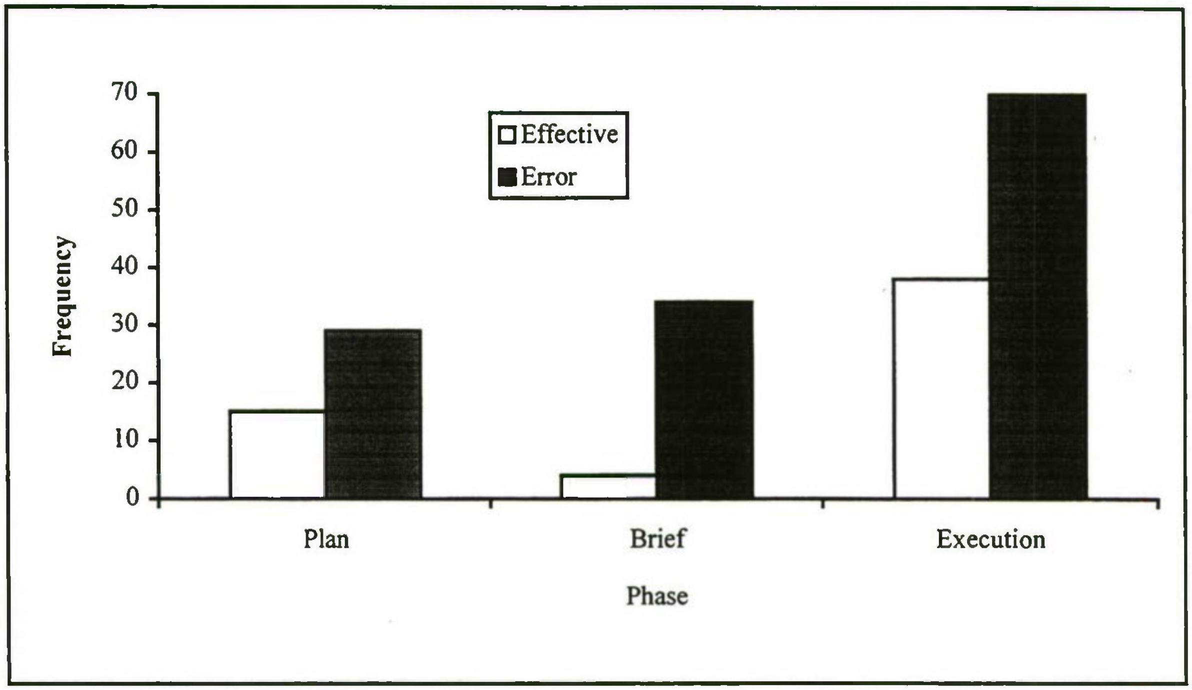


Figure 3. Frequency of effective communication behaviors and errors across phases.

Finally, Figure 4 shows how these behaviors have been further delineated into specific breakdowns in communication across the mission phases. From these detailed data we are beginning to build and validate a model of large distributed team performance and to understand the relationships among behaviors.

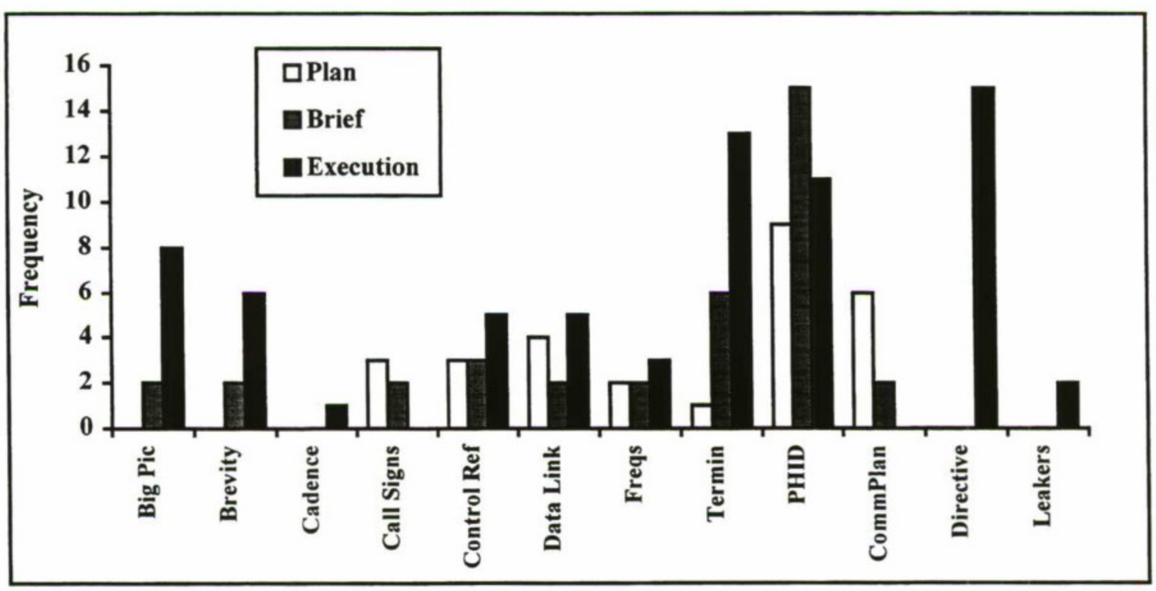


Figure 4. Specific breakdowns in communication across mission phases.

Post-Detachment Training Utility Questionnaire

We found that several of the items in the training utility questionnaire were highly correlated. Thus, it might be possible to collapse them into a smaller number of constructs. Principal components extraction with varimax rotation was performed on 22 of the 24

questionnaire items (two factors were dropped due to their utility as outcome measures). The five factors which were extracted from the set of 101 participants are presented in Table 2.

Table 2. PRELIMINARY FACTOR ANALYSIS

Mental Models	Implementation	Orientation	Adaptation	Individual Skills
Mission Types	Tactical Skills	Planning Missions	Contingency Planning	Pre-Trng Prep
Own A/C Role	Decision Making	Briefing Missions	Threat Awareness	Flight Discipline
Asset Relationships	Situation Awareness	Integration/ Coordination		Personnel Familiarity
Mission Chars	Comms			
Other Plats Capabilities	Mission Debrief			
Other Squadron Procedures				

These five factors were then used in a series on MANOVAs to determine the extent to which demographic data predicted factor loadings. Participants were divided into three groups based on experience with carrier operations: those who had never been deployed, those who had been on one deployment, and those who had been on two or more deployments. A Hotteling's multivariate test of significance achieved an approximate F=2.817; $p<.01$. These results suggest that aviators with differing levels of experience have different loadings on the five factors. The univariate F tests show that both the mental model factor [$F(2,96) = 5.408$; $p<.01$] and the adaptation factor [$F(2,96) = 4.182$; $p<.01$] were significant. These data, depicted in Figure 5, suggest that the utility of training for mental model development decreases as experience increases. However, as an aviator's experience increases, their ratings of training utility suggest that they are learning more about how to adapt to the dynamic environment in which they operate.

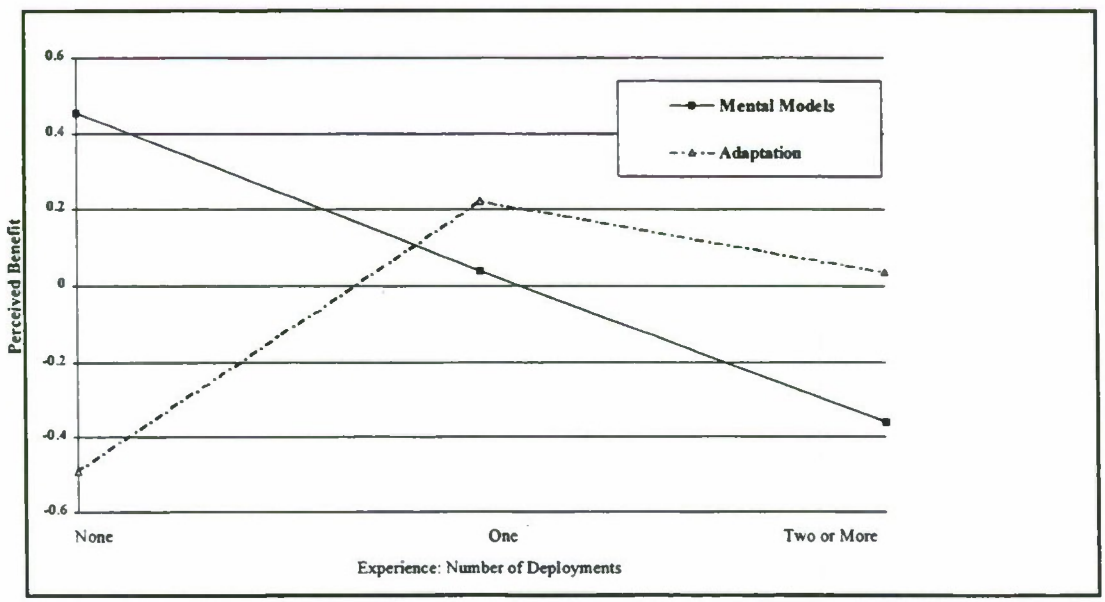


Figure 5. Utility of training for mental model development.

DISCUSSION

The behavioral data collected to date have begun to highlight specific areas in the planning, briefing, and execution phases of the mission that may benefit from additional training. The preliminary analyses of the behavioral observations suggest that for large, distributed, tactical teams with limited on-line communication opportunities, tactical performance is related to the planning and briefing. We are in the process of building and validating a model of the relationships between specific behaviors across phases. Planning included the effective use of assets, coordinated tactics development, specified communications priorities, and detailed contingency planning. In order to effectively plan, it appears that the planning team must have an in-depth knowledge of types of missions that might accomplish the objectives, the combination of assets that can/should be used given specific sets of circumstances (e.g., weather, target, threats), the coordination requirements of the mission given the assets employed, and an understanding of what can/might go wrong.

Briefing was the communication of the plan to the rest of the team and included the information presented, the order of presentation, and the how the information was presented (e.g., tactical timelines, general and specific tactical pictures). During this phase it was crucial to convey the "Big Picture" to provide the aviators with a framework within which to understand how their performance fit in with the team objectives. The briefing has the potential to set the stage for SA within execution by providing the team with a mental model of how the mission should progress.

Good briefings included a clear presentation of the objectives, the roles and responsibilities of each subteam, a clear timeline of what should be happening and when, picture of the what whole team should be doing where at several critical points throughout the mission, and, finally, provide contingency plans that lay out specific conditions for changing the mission the in flight and the specific plans for what would then be the new tactical requirements. These aspects of good briefing support the notion that briefs help the aviators form a mental model of the upcoming mission. By providing a template of expected performance and potential contingencies, the aviators are better able to monitor the performance of the team, as well as their own performance, by focusing their attention on specific cues and patterns in the tactical environment. This may improve their ability to execute precisely coordinated missions with minimal explicit coordinated missions with minimal explicit coordination (i.e., communication).

The results of the training utility data provided support for the notion that novices might benefit from basic training in air wing operations prior to these large team training events. While we had a relatively small sample of training utility respondents (n=101), significant differences were found between novices and experts in the perceived training utility of team training events. The items targeting the development of mental models addressed knowledge of team tasks, roles, capabilities, responsibilities, and interactions (Cannon-Bowers, Converse, & Salas, 1993). It is of particular interest that the utility of training for novices was highest for the development of mental models, while experts perceived more value for developing skill for adapting to the dynamic combat environment (Figure 5). This suggests that the novices are focusing on building up basic models of the mission, roles and relationships, while the experts who already have this basic knowledge are now able to focus on developing strategic adaptation skills, such as contingency planning/execution and threat awareness.

These findings led to a modification of our conceptual model describing how planning and briefing might affect the performance of large tactical teams. In Figure 6, we have added the nature of the large team that must use the briefing as their primary source of mission information as a variable that interacts with the brief to affect team performance. Given the limited timeframe in which this information must be conveyed, it would seem important that the team members have a well developed cognitive framework in which to assimilate the mission specifics. However, it is not uncommon for an air wing team to have up to 50% novices (no carrier deployments or coordinated air wing operations). One might expect that the ability of these junior aviators to quickly and completely understand the mission requirements and coordination demands might be less than optimal. One might also expect that this lack of a complete mental model of the team

task might negatively impact their ability to maintain awareness of where they are and where the team is in relation to the plan.

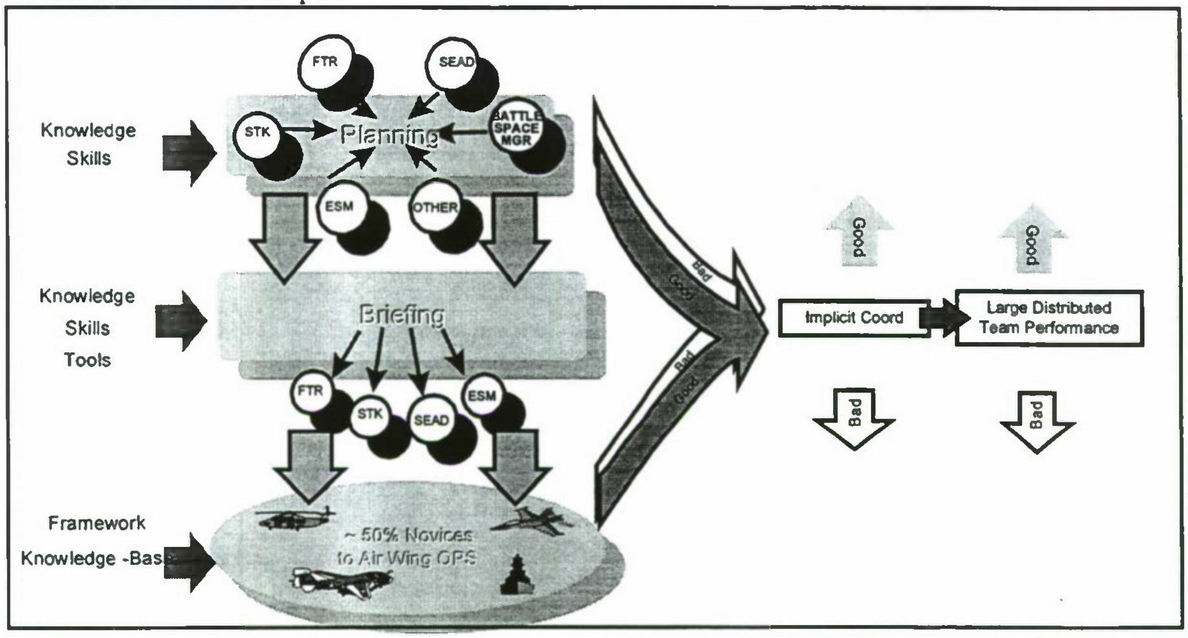


Figure 6. Modified conceptual model of factors affecting large distributed team performance.

Methods to facilitate the integration of large tactical teams continued to shape this research program. The question of what knowledge representations are critical for each air wing element to successfully plan, understand and execute missions, as well as the question of the nature of the information that should be shared among team members present our primary challenge. The goal in pursuing these questions is to develop training strategies and guidelines for enhancing the readiness of air wings and to improve planning and briefing strategies and technologies.

REFERENCES

Cannon-Bowers, J. & Salas, E. (1990). Cognitive psychology and team training: Shared mental models in complex systems. Paper presented at the 5th annual meeting of the Society for Industrial and Organizational Psychology, Miami, Florida.

Cannon-Bowers, J. A., Salas, E., & Converse, S. (1993). Shared mental models in expert team decision making. In J. Castellan Jr. (Ed.) Current issues in individual and group decision making (pp. 221-246). Hillsdale, NJ: Erlbaum.

Salas, E., Prince, C., Baker, D. P., & Shrestha, L. (1995). Situation awareness in team performance: Implications for measurement and training. Human Factors, 37, 123-136.

Stout, R. J. (1995). Planning effects on communication strategies: A shared mental model perspective. Proceedings of the Human Factors & Ergonomics Society 39th Annual Meeting, Santa Monica, CA, 1278-1282.

This page intentionally blank.

ENHANCING SITUATIONAL AWARENESS WITH ABOVE REAL-TIME TRAINING (ARTT)

Peter Crane
Air Force Research Laboratory
Aircrew Training Research Division,
Mesa, AZ

Dutch Guckenberger
ECC International Corporation /
University of Central Florida,
Orlando, FL

In the early 1970s, engineers and test pilots at the National Aeronautics and Space Administration's (NASA) Dryden Flight Research Center apparently increased the effectiveness of flight training simulators by deliberately distorting simulated time. Kolf (1973) and Hoey (1976) briefly document simulator training interventions which were aimed at improving test pilots' ability to keep up with the pace of events in flight. Kolf notes that, "regardless of the type or amount of pre-flight simulator training accomplished by the pilot, the actual flight seems to take place in a much faster time frame than real time," (p. 1). Hoey (1976) reports that in the X-15 program, pilots typically spent ten hours in the simulator for each ten minutes of flight. Even with this preparation, pilots reported that, "It sure seems to happen faster in the real airplane," or, "I had the feeling that I was 'behind the airplane'," (pp. 2-3). As an experiment, Kolf increased the rate of simulated time in the M2-F3 Lifting Body simulator. In the modified simulator, a mission profile which normally required 10 minutes to complete took place in only 6 minutes, 40 seconds. Three experienced M2-F3 pilots flew a familiar mission at 1.5 times real-time and all agreed with "enthusiastic responses," (p. 2) that the modified simulator felt exactly like the aircraft. A second application of fast-time simulation (Hoey, 1976) was to a flight test program for remotely piloted vehicles (RPV). RPV pilots who used simulation at 1.4 times real-time as final preparation before a flight reported being, "Less rushed and more confident," (p.18) than when using real-time training exclusively.

The NASA application of Above Real-Time Training (ARTT) was limited to training expert pilots preparing for specific missions. Manipulating apparent time has been evaluated more recently as an instructional tool for both novices and experienced individuals. Schneider, Vidulich, & Yeh (1982) and Vidulich, Yeh, & Schneider (1983) used time-compression to help train air traffic controllers. The task for these controllers was to monitor an aircraft's flight path on a radar display and issue turning instructions so that the aircraft would fly through a specific vector. Actual aircraft would traverse 20 nautical miles and require approximately five minutes at 260 knots to complete the turn. These researchers increased the apparent rate of time in the simulator to 20 times real-time so that a turn would be complete in approximately 15 seconds. Vidulich, Yeh, & Schneider (1983) trained university students over four hours to perform a turn point task. A group of students who performed the task in real-time experienced approximately 32 trials in four hours of training. A group performing the same task using ARTT at 20 times real-time received approximately 260 time-compressed trials followed by only 3 or 4 real-time trials in four hours of training. All trainees were tested at real-time for two hours. ARTT subjects showed significantly better performance at initiating turns properly. These authors assert that the ARTT improves training effectiveness by allowing many trials and training under a mild speed stress.

Guckenberger, Uliano, & Lane (1992) trained novices, university students, in tank gunnery using several ARTT conditions. In this experiment, students were trained in gunnery tasks which required them to detect, identify, and shoot a moving target using an M1 tank part-task trainer. Students received five familiarization trials in real time followed by 15 training trials in real time or in one of four ARTT conditions. Students were then tested in real time. Subjects in all four ARTT groups showed better performance on test trials than the students trained in real time. Guckenberger, Uliano, Lane, & Stanney (1992) conducted an experiment using 24 experienced F-16C pilots. Pilots trained at real-time, 1.5 times real time, 2.0 times real-time, or with a random mix of apparent times. Pilots then tested at real-time. One task required the pilot to engage a bandit and to perform a complex threat response when a warning was detected. For this third, dual-threat task, the 2.0 times real-time and mixed ARTT groups showed faster threat response than the group trained in real-time and all ARTT groups achieved significantly more bandit kills during real-time, test trials.

Schneider (1989) proposes that the primary effect of time compression is to allow more training trials within a given period of clock time. In the air traffic control studies, subjects were given the same

amount of training time in the simulator so that the ARTT subjects received more training trials. In contrast, Guckenberger et al. gave all subjects the same number of training trials so that the ARTT subjects received less training time than the students trained in real time. Since the students trained using ARTT performed better on real-time test trials than students trained in real time, Guckenberger et al's results indicate that ARTT has a beneficial effect beyond simply increasing the number of training events.

The present research effort has focused on application of ARTT to air radar interpretation/air intercept and emergency procedures. Previous research on ARTT has used university students as trainees and/or lower fidelity simulators. This research employed Air Force F-16 pilots and student pilots, high-fidelity simulators, and training problems which emphasize skills required for air combat.

Situation Awareness Context. Researchers studying situation awareness have stressed the need for pilots to rapidly encode and understand the significant events occurring in the environment (Sarter & Woods, 1991). Researchers studying expertise have found that these are the abilities which characterize expert level performance. Further, research has demonstrated that expertise is based on rapid access to an extensive knowledge base of domain specific information. The expert identifies the present situation as an example of a familiar prototype and selects a response accordingly. Fracker (1988) asserts that an identical mechanism is responsible for situation awareness: "Matched knowledge structures.., provide the pilot's assessment of the situation and serve to guide his attention," (p. 102). The authors assert that ARTT may be a methodology that aids pilots in rapidly encoding and understanding the significant events occurring in the environment and by implication enhance SA.

EXPERIMENT 1: ARTT FOR EMERGENCY PROCEDURES TRAINING WITH EXPERIENCED PILOTS

For this experiment, experienced F-16 pilots conducted single-ship, defensive counter-air missions over a ground target using two scenarios. In one scenario, single-emergency, the pilot's aircraft suffered engine failure. The pilot's task was to re-start the engine and then to engage an incoming bandit. In the second scenario, multiple-emergencies, the pilot had to respond to several emergencies and engage two bandits in succession. Pilots received initial training in real time followed by additional practice in real-time or at 1.5 times real time. All pilots were then tested in real-time.

Research Methods

Participants. The participants in this experiment were 12 F-16C pilots from the 347th Fighter Wing, Moody Air Force Base, Georgia. Pilot experience ranged from 150 to 1600 F-16 hours.

Apparatus. An F-16 trainer developed by the ECC International Corporation, Orlando, FL, was selected for this experiment. The ECC F-16 simulator was developed for the Air Force Unit Training Device program and later modified for ARTT research. The ECC simulator incorporates F-16 aerodynamics and avionics capabilities with a three-screen, out-the-window visual display system. The system has the capability to present scenarios in which other aircraft fly in pre-recorded flight paths. This system was modified for ARTT by altering the software time integration factor. The amount of assigned simulated time between frame updates was increased by a factor of 1.5. The simulation model produces the same number of frames updates as during real-time operation but more simulated time has passed between updates.

Procedure. Pilots were randomly assigned to either the RTT or ARTT condition. All pilots received familiarization training in the ECC F-16 simulator in real time. Familiarization consisted of flying vertical-S maneuvers, 90° turns, and loops. After familiarization, pilots were trained and tested in one of the two emergency conditions (single or multiple) selected at random. After a break, the pilot was trained and tested in the other condition. For both conditions, pilots received initial training in real-time until they could complete the task without failures or re-starts. Typically, initial training required three or four trials. Data from these trials were not analyzed. Six additional practice trials were then conducted in real time or above real time. Finally, all pilots received four, real-time, test trials using the same scenarios as used in training. The dependent measures were time required to correct emergencies and time required to kill the bandits.

Single emergency. Pilots were initialized at 2000' AGL and 480 knots inbound toward a power plant which they were tasked to defend against air assault. Shortly after the trial started, the F-16's engine failed and the pilot had to restart using established procedures. Immediately after restart, the pilot engaged a MiG-29. The trial ended when the bandit aircraft was killed. Dependent variables were time to restart engine and time to kill the bandit.

Multiple emergency. The multiple emergency task was similar to the single emergency except the pilot first had to check an equipment hot indicator light, restart a failed engine, acknowledge another indicator light (hydraulic failure), and then engage a MiG-29 followed by a second MiG-29. Dependent variables were time to respond to the equipment hot light, time to restart the engine, time to respond to the hydraulic failure indicator light, and time to kill both bandits.

Results

Single emergency. Mean time to restart the engine and time required to kill the bandit on test trials are plotted on Figure 1.

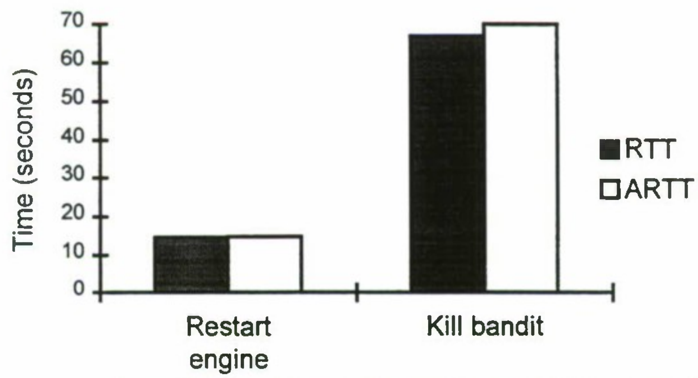


Figure 3. Mean Time to Restart Engine Kill Bandit for Test Trials.

Test performance on both variables was not affected by training condition. Time required to restart the failed engine, $t(11) = 0.08$, *ns*, and time required to kill the bandit, $t(11) = 0.4$, *ns*, were not significantly different between the RTT and ARTT pilots.

Multiple emergency. Mean time on test trials to respond to the equipment hot light, time to restart the failed engine, time to respond to the hydraulic failure light, and time to kill both bandits are plotted on Figure 2.

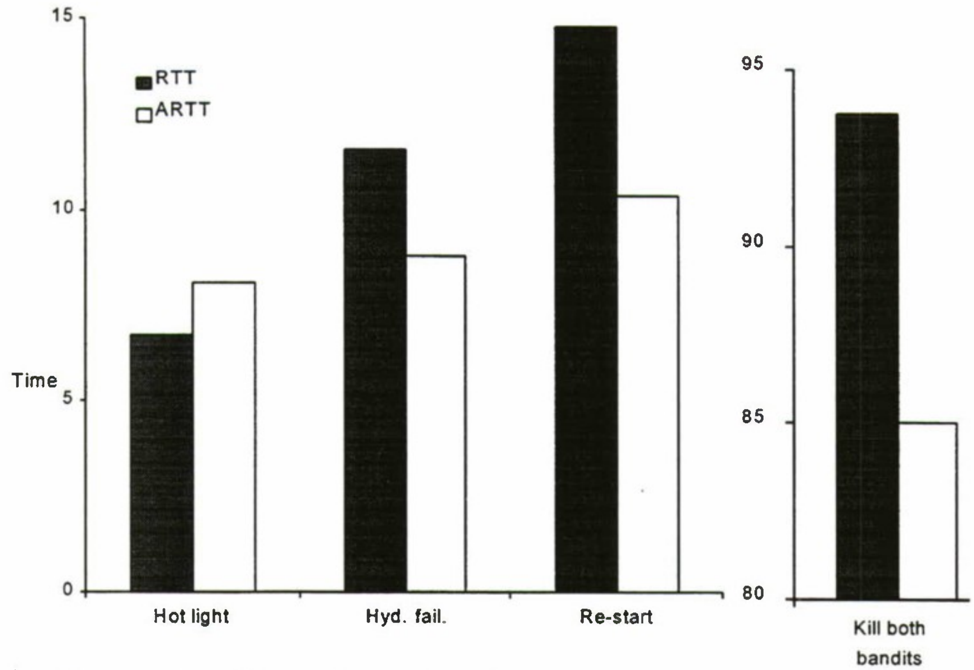


Figure 4. Mean Time to Respond to Hydraulic Failure Test Trials.

On real-time test trials, time to clear the equipment hot light was not significantly affected by training condition, $t(11) = 1.65$, *ns*. However, times required to restart the failed engine, $t(11) = -4.54, p < .001$, respond to hydraulic failure, $t(11) = -4.93, p < .001$, and kill both bandits, $t(11) = -2.76, p = .018$, were all significantly less for pilots trained using ARTT than for pilots trained using RTT.

Discussion. ARTT pilots in this experiment replicated the training procedures used by NASA in that the pilots first experienced initial training in real-time followed by additional trials at 1.5 times real-time. The RTT pilots received all training trials in real-time. The experience of NASA pilots was that ARTT provided better preparation for highly demanding missions than real-time simulation. In this experiment, ARTT was more time efficient than RTT for the single emergency task but provided no additional training benefits. Pilots in the single emergency task performed engine restarts and defeated a bandit aircraft as quickly as pilots trained in real-time but no faster. However, in the more complex multiple emergency task, pilots who practiced using ARTT were able to perform two of the three required emergency procedures and, killed both bandit aircraft significantly faster than pilots who trained in real time. For these experienced F-16 pilots, the single emergency task which consisted of restarting the engine and engaging a single bandit was not especially demanding. ARTT provided no training advantage for this task other than reducing the amount of clock-time required to complete a given number of practice scenarios. The multiple emergency task was more demanding of the pilot's time and workload management skills. For this task, practice using ARTT after initial training in real-time helped the pilot to perform the emergency procedures and to successfully engage both bandits significantly faster than pilots who received all of their training in real time.

Conclusions. These results support the hypothesis that ARTT provides improved training for some tasks compared to conventional, real-time training. Pilots trained using ARTT performed responses in the single emergency task as well as pilots trained in real-time. ARTT was more efficient than real-time training in that pilots trained using ARTT were able to perform on test trials as well as pilots trained in real time but with less training time. For the more demanding multiple emergency task, pilots trained using ARTT performed emergency responses faster than pilots trained using RTT for three of the four dependent variables. ARTT provided better training than real-time training provided that: a) the tasks being trained are highly demanding of a pilot's time and workload management skills, and b) the pilot has received initial training in real-time. ARTT provided stressed training and pilots were required to develop efficient time and workload management strategies. These strategies were reflected in faster responses during test

trials. By implication, the ARTT trained pilots were able to perform more tasks in a given amount of time which should lead to increased SA.

EXPERIMENT 2: ARTT RADAR SKILLS TRAINING WITH STUDENT PILOTS

In experiment 1, the trainees were mission-ready fighter pilots who were well trained in the tasks that were simulated but lacked recent experience. For these trainees, ARTT produced equal or better test performance with less training time than real-time simulation. In experiment 2, the training benefits of ARTT were assessed with student pilots in a Formal Training Unit (FTU). In this experiment, a radar-skills task was used as a supplement to an existing training syllabus. Students practiced radar skills within a mission context after they had successfully completed the air-to-air portion of the F-16 FTU syllabus. Further, in experiment 1, all pilots received the same number of training trials with the ARTT pilots receiving less time (clock hours) in the simulator than the pilots trained in real-time. In experiment 2, ARTT pilots received more training trials using approximately the same amount of clock-time in the simulator than the pilots trained in real time. In this respect, experiment 2 replicated the procedure used by Vidulich, Yeh, and Schneider (1983) who used ARTT to provide more training in a given time period than could be provided using real-time simulation.

Research Methods

Participants. The participants in this study were 24 students in the F-16C training course, 58th Fighter Wing at Luke Air Force Base, Arizona. All participants were new to the F-16 with between 40 and 130 F-16 hours. Of the 24 pilots, 19 had no previous Air Force flying experience other than Undergraduate Pilot Training and Lead-in Fighter Training for a total of 260 - 615 flight hours. The remaining pilots had previous assignments in other aircraft which were not equipped with air-to-air radar. These pilots had 1500 - 2100 hours in other aircraft but only 50 - 100 F-16 hours. All pilots had completed the air-to-air portion of training and had successfully completed simulator and aircraft sorties requiring use of the air-to-air radar.

Apparatus. The Armstrong Laboratory Air Intercept Trainer-Plus (AIT+) was selected for this experiment. The AIT+ is an Armstrong Laboratory F-16 Air Intercept Trainer which has been modified by replacing the computing hardware and software with components from the Armstrong Laboratory Multi-Task Trainer (Boyle and Edwards, 1992). The AIT+ is a high-fidelity F-16C simulator limited to air-to-air operations. The AIT+ incorporates flight, engine, and radar simulations, with hands-on-throttle-and-stick (HOTAS) controls, a radar display, radar control panel, and a color monitor which includes a heads-up display (HUD) and a limited out-the-window display. For this experiment, the AIT+ operated in autopilot mode in that the aircraft's altitude, airspeed, and heading were fixed. The AIT+ was modified for ARTT by changing the time integration factor.

Procedure. Pilots were randomly assigned to the RTT or the ARTT group. The objective of radar skills training was to increase pilot proficiency in using air radar to search and sort multiple, maneuvering targets. For the radar skills task, pilots received the following instructions:

> For the radar skills task, your F-16 will fly on auto-pilot on a straight and level course into a bandit formation. All bandits in the scenario will fly preplanned routes. Your task is to use the radar as effectively and efficiently as possible to provide you with all of the critical information concerning the inbound bandits. When performing the radar skills tasks, your performance and the evaluator's grading will depend on your ability to:

- Search all airspace (surface to 50k) before the closest bandit is within 40nm.
- Know the initial picture (number of bandits, formation, aspect, altitude, airspeed).
- Determine bandit actions (maneuvers).
- Know the picture after bandit actions.
- Know which bandit is the highest threat.

The pilot's F-16 was initialized at 15,000 feet, 450 knots airspeed, and heading 360.

The radar skills task began with a relatively simple scenario (Figure 3). The pilot was to call out radar contacts and bandit actions as they occurred. The evaluator, a retired F-16 instructor pilot and squadron commander, stood just behind the cockpit and could view the radar screen on a video monitor.

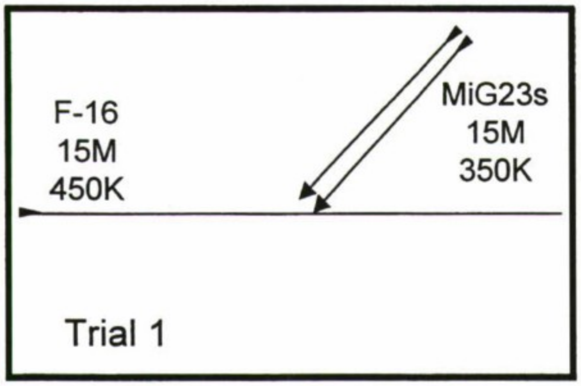


Figure 3. Simple Scenario from the Radar Skills Task.

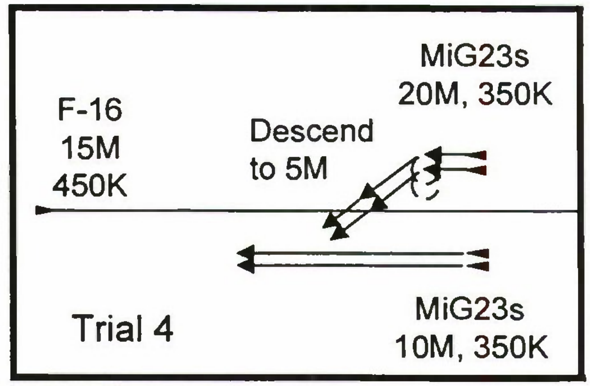


Figure 4. Moderately Complex Scenario from the Radar Skills Task.

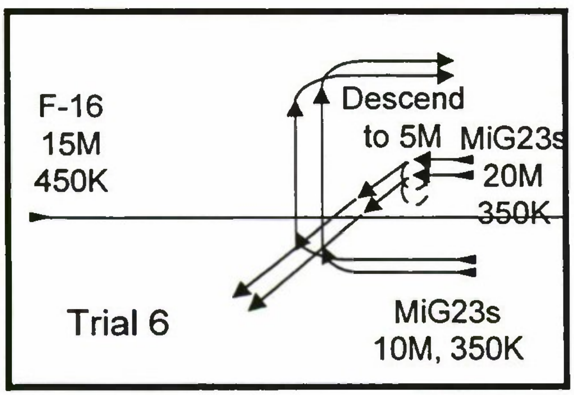


Figure 5. Difficult Scenario from the Radar Skills Task.

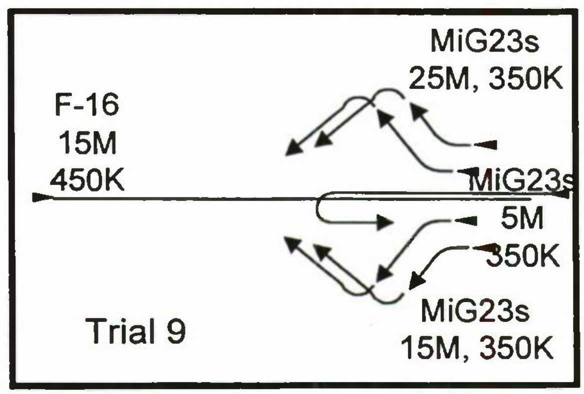


Figure 6. Test Scenario (most difficult) from the Radar Skills Task.

The pilot's radar skills performance was scored on a scale of 0 - 3 for each of four sub-tasks: search the airspace and sample the contacts; sort the formations and monitor actions; describe picture after bandit action; and target the highest priority threat before coming within 10nm. The pilot's scores for each of the four sub-tasks were summed for a composite run-time score. After passing the bandit formation, the simulation stopped and the evaluator asked the pilot to debrief the scenario. Debrief consisted of four sub-tasks: describe the initial picture, bandit actions, the picture after bandit action, and the factors used to determine the highest threat. Debrief performance was scored using the same scale as for the run-time scores. During training trials, the evaluator provided feedback on the scenario after the pilot had completed his debrief.

Pilots received 10 or 15 training scenarios. Pilots in the RTT condition received two relatively simple scenarios, three moderately complex scenarios, and five complex scenarios. Pilots in the ARTT condition received the same simple, moderate, and complex scenarios plus five additional complex scenarios for a total of 15 training trials. All pilots were tested in real-time on five scenarios which were more complex than any of the training scenarios.

Results. Scores for training and test trials were grouped into blocks depending on scenario difficulty. Trials 1 - 5 were grouped as simple - moderate complexity scenarios, trials 6 - 10 as difficult, and trials 11 - 15 also as difficult but for the ARTT group only. Trials 16 - 20 were test trials and were designed to be more difficult than any of the training trials. Mean percent scores grouped into blocks are

plotted for run-time scores on Figure 7 and for debrief scores on Figure 8. Test performance was not significantly different between the RTT and ARTT groups for run-time scores ($\overline{X}_{RTT}$ = *77.5*, $\overline{X}_{ARTT}$ = *81.7*, *t(22) = 1.97, p = .062*). For debrief scores, test performance for the ARTT group was significantly higher than for the RTT group ($\overline{X}_{RTT}$ = *75.8*, $\overline{X}_{ARTT}$ = *82.5*, *t(22) = 2.38, p = .026*).

Comparing scores on training trials and test trials, run-time scores for the ARTT group show significant increase from training to test (*F(1, 11) = 19.64, p = .001*) while for the RTT group there is no significant change in run-time scores from training to test (*F(1, 11) = 1.44, p = .256*), (Figure 7). For debrief scores, the ARTT group shows a significant increase from training to test (*F(1, 11) = 10.76, p = .007*) while the RTT group shows a significant decrease (*F(1, 11) = 38.71, p < .000*), (Figure 8).

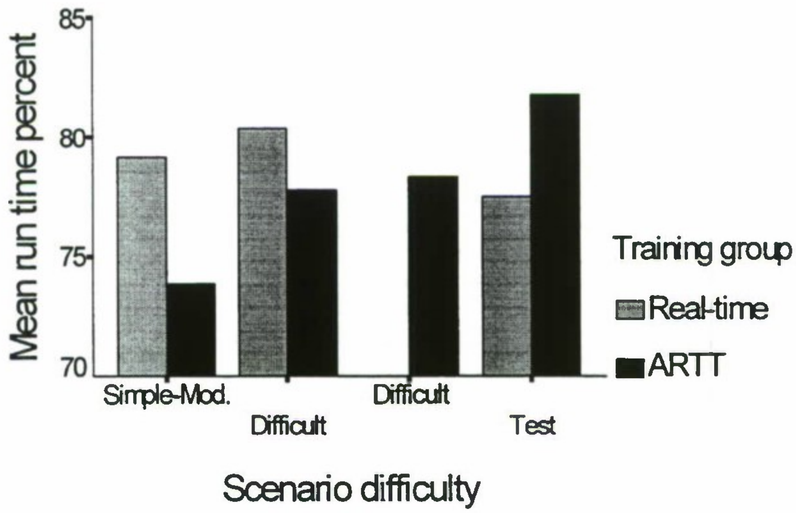


Figure 7. Run-time Scores Grouped by Scenario Difficulty.

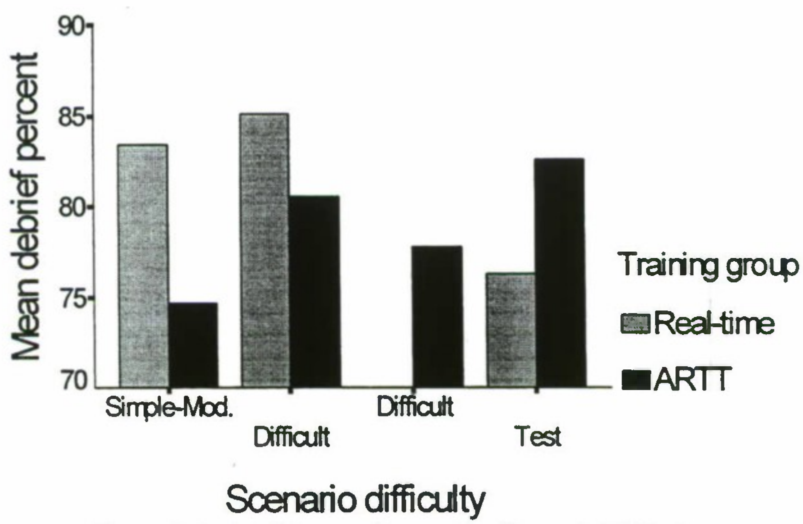


Figure 8. Debrief Scores Grouped by Scenario Difficulty.

Discussion. ARTT will interfere with effective training if the pace of events is so rapid that students cannot keep up. Evidence for this effect would be that students using ARTT would produce lower performance scores than students using RTT and that ARTT student performance would not improve. In this experiment, performance scores of students using ARTT were indeed initially lower than scores of students using RTT, however, ARTT student performance significantly improved from trials 1-5 (simple-moderate scenarios) to trials 6-10 (difficult scenarios) for run-time but not for debrief scores. Performance

scores for students trained in real-time did not change significantly from the simple-moderate scenarios to the difficult scenarios. Overall, ARTT at 1.5 times real-time was not a problem for advanced student pilots although there was an initial performance deficit. ARTT did require the students to keep up a rapid pace but they learned to do so after the first five training trials. What was important was that ARTT allowed the students to develop a more extensive knowledge base regarding multi-bandit radar intercepts. The increased knowledge base supported improved performance on the more complex test trials. ARTT enhanced SA for students by increasing the time efficiency of simulator training. In sum, ARTT aids development of SA by providing stressed training on demanding tasks for experts and more time efficient training for students.

In experiment 1, ARTT provided equal or improved training using less time spent in the simulator. In experiment 2, ARTT pilots received more training trials than the RTT pilots while time spent in the simulator was approximately equal for both groups. Student pilot debrief scores on the test trials were significantly higher for the ARTT group than for the RTT group. Further, the ARTT group showed a significant increase in run-time and debrief scores from the difficult training trials to the more complex test trials which were presented in real time. The RTT group showed no change in performance between training and test trials for run-time scores (Figure 7) and a significant decrease for debrief scores (Figure 8). Overall, the idea of using ARTT to provide additional training trials within the same amount of clock time as real-time training as suggested by Vidulich et al. (1983) was supported.

SUMMARY AND CONCLUSIONS

The concept of above real-time training was developed by engineers and pilots as a practical solution to an immediate problem. There are two proposed advantages of ARTT. The first is simple efficiency. Using time-compressed simulation, a pilot using ARTT can experience a given number of training events in fewer clock-hours of simulator time or, the pilot using ARTT could experience more training events in a fixed amount of training time. The second proposed advantage to ARTT is that ARTT should provide easier transition to the more demanding environment of actual flight than normal-time simulation. The experience of NASA test pilots was that actual flight was more demanding than simulation; ARTT felt more like the airplane than a high-fidelity, real-time simulation.

In the first experiment, experienced F-16 pilots performed air-combat tasks while responding to in-flight emergencies. Pilots trained using ARTT performed a single emergency procedure and engaged a single bandit as well as pilots trained in real-time but no better. ARTT was more efficient than real-time training in that pilots trained using ARTT were able to perform on test trials as well as pilots trained in real time but with fewer clock-hours of simulator time. However, in a more demanding multiple emergency task, pilots who received initial training in real time followed by additional practice using ARTT performed emergency procedures and engaged two bandits faster than pilots who received all training in real time. This experiment supports the hypothesis that ARTT can provide better transition to a more demanding task environment than real-time training for selected tasks.

In the second experiment, student F-16 pilots performed a radar skills task. Unlike the more experienced pilots in experiment 1, performance of student pilots on this task was initially degraded by ARTT, however, ARTT pilot performance improved with additional training trials. In this experiment, ARTT was used to increase the number of training trials presented within approximately 30 minutes of simulator time. Student pilots trained in real-time received ten training trials while pilots trained using ARTT received fifteen trials. Pilots trained using ARTT performed better on real-time, test trials than pilots trained in real-time. The combination of ARTT plus additional training trials led to improved performance on the test trials without increasing clock-hours in the simulator.

As an instructional strategy, ARTT is inexpensive to implement and can increase time efficiency and training effectiveness for many training tasks. The hypothesis that ARTT is more efficient than real-time simulation was supported. Compared to real-time simulator training, time-compressed training that does not overload the trainee can support equivalent levels of test performance with fewer hours of clock-time in a simulator or improved test performance with equal amounts of clock-time in a simulator. The hypothesis that transition to a more demanding task will be increased by ARTT without additional training trials was also supported for some tasks. Pilots who received initial training in real time followed by additional practice using ARTT performed faster than pilots trained in real time for highly demanding

tasks. The initial results for ARTT are encouraging, and the authors suggest that the potential for ARTT to enhance SA warrants further research.

REFERENCES

Boyle, G. H., & Edwards, B. (1992). Low cost trainers: Lessons for the future. *Proceedings of the 14th Industry/Interservice Training Systems and Education Conference*, Orlando, FL: National Security Industrial Association.

Fracker, M. L. (1988). A theory of situation assessment: Implications for measuring situation awareness. Proceedings of the Human Factors Society 32nd Annual Meeting, Santa Monica, CA: The Human Factors Society.

Guckenberger, D., Uliano, K., & Lane, N. (1992) The application of above real-time training (ARTT) for simulators: Acquiring high performance skills. *Proceedings of the 14th Industry/Interservice Training Systems and Education Conference*, Orlando, FL: National Security Industrial Association.

Guckenberger, D., Uliano, K., Lane, N., & Stanney, K. (1993). The effects of Above Real-Time Training (ARTT) on three tasks in an F-16 part-task simulator. *Proceedings of the 15th Industry/Interservice Training Systems and Education Conference*, Orlando, FL: National Security Industrial Association.

Hoey R. G. (1976). Time compression as a means for improving the value of training simulators. Unpublished manuscript.

Kolf, J. (1973). Documentation of a simulator study of an altered time base. Unpublished manuscript.

Schneider, W. (1989). Getting smarter quicker: Training more skills in less time. *Science and Public Policy Seminar*, Washington, DC: Federation of Behavioral, Psychological, and Cognitive Sciences.

Schneider, W., Vidulich, M. & Yeh, Y. (1982). Training spatial skills for air-traffic control. *Proceedings of the Human Factors Society 26th Annual Meeting*. Santa Monica, CA: Human Factors Society.

Vidulich, M., Yeh, Y., & Schneider, W. (1983). Time compressed components for air intercept control skills. *Proceedings of the Human Factors Society 27th Annual Meeting*. Santa Monica, CA: Human Factors Society.

This page intentionally blank.

MISCELLANEOUS

This page intentionally blank.

EMERGENT DISPLAYS: TECHNOLOGY AND ISSUES IN SITUATIONAL AWARENESS

Lisa B. Achille
Human Computer Interaction Laboratory
Naval Research Laboratory*
Washington, DC

Daniel L. Dolgin
Naval Aviation Science and Technology Office (NAVSTO)
Arlington, Virginia

Terry Allard
Cognitive and Neural Science and Technology Division
Office of Naval Research
Arlington, Virginia

The ability of the pilot to maintain situational awareness (SA) is recognized in the aerospace community as critical to mission safety, survival, and completion. Establishing and maintaining aircrew SA must become a predominant design goal in developing future aircraft display systems. While accomplishments in the scientific community to define, explore, and measure this significant cognitive construct have allowed cockpit design and display efforts to realize numerous goals, the need exists for continued efforts at improving human performance. This paper discusses critical issues and areas for future research in regard to cockpit display designs and human performance

Introduction

Better information management leads to better SA. Cockpit display technology can either improve SA or overload the pilot with too much information. Increasing information density in the cockpit has radically changed the pilots information processing task from that required by early aircraft displays. Today's cockpit tasks have evolved from relatively simple "stick-and-rudder" control to complex command center decision-making. Cognitive, attentional/perceptual resources, and psychomotor control are often pushed to their limit, due to the introduction of numerous independently developed information and control systems. The challenge of today's aircraft designers is the integration and management of increasingly complex sensor, weapons, and control systems supporting today's information warriors.

This trend toward increased information density will continue into the 21st century as Navy aircraft become engaged in joint operations in the compressed littoral battle space (AGARD, 1997). Advances in the capabilities of sensors and data fusion, combined with evolving battlefield communications will radically change our vision of SA and magnify the importance of effective and efficient human-machine systems design. Target detection capabilities will dramatically improve by the year 2020 through the use of Moving Target Indication and Automated Target Recognition technologies. New systems will be developed to track friendly forces and evaluate their status at more detailed levels than is currently possible, dramatically reducing the need for situation report requests. Heads-up displays, glass cockpit displays, spatialized audio, and targeting by eye tracking are examples of the rapidly changing and emerging cockpit display and control technologies that will provide additional sources and types of information to pilots. The potential exists for even the most experienced aviator to be overwhelmed with information processing difficulties under normal flight conditions. The essential paradox is that information intended to help pilots could degrade performance by overloading their information processing ability/capacity. As new technologies are developed and incorporated into cockpit systems, the balance of workload between warfighters and the systems that aid them will become an increasingly important design factor.

Challenges to cockpit display design, highlighted over fifteen years ago, still remain true today. The pilot is required to maintain optimum SA throughout a mission, even under high stress and high workload situations. The capability of displaying more information in less information panel space does not, however, translate directly to reduction in crew workload. "The reduction of crew workload through the application of airborne computers and electronic displays must be achieved through the efficient display of essential, nonsuperfluous information (McGee, 1982)." "One major problem with electronic innovation is that combat pilots are now supplied with more data than they can effectively use (Aronson, 1981)." These comments remain true, in part, because the tools and knowledge needed to design display systems for human use and to enhance SA have lagged behind the stunning advances made in display technology.

The introduction of new technologies and computer-assisted systems in the cockpit will not necessarily overcome operator workload problems. In fact, the addition of new technologies, e.g., increasingly complex multi-modal visual displays, may have a detrimental effect on pilot safety due to

distractibility, especially under stress. For example, Rockwell, Giffin, and Romer (1983) found that pilots became so engrossed in their attempts at diagnosis they paid no attention to their flight position. This behavior resulted in the majority of the pilots violating their altitude clearance. Hardware development that is conducted independently from human-systems design is quite likely to require significant and costly modification or even failure. Bost has ascribed this result to "technology myopia," the "deep-seated belief that if the technology is done correctly, then all will be OK (Bost, 1996)." The use of new and costly display hardware is no guarantee the human-machine system will operate correctly. The management of the displayed information is key to optimal human-machine performance under all workload conditions.

Pilots get more of their information today from cockpit displays than "out-the-window" views of the environment when compared to the aircraft of 20 years ago. Technological advances and increasing task demands continue to reduce the pilots' sensory contact with the environment distancing them from the physical environment outside the vehicle. The pilot must integrate information about the current situation from multiple instruments, sensors, and computer generated images. Therefore, the pilots SA can be helped or hindered by the design of the cockpit systems. If effective SA can be enhanced through competent system design, we will significantly enhance pilot performance and decrease the possibility of expensive and tragic accidents.

Situational Awareness (SA) Defined

The need to relate function to design has become even more important with the pilot's changing role in the battlespace of the 21st century (Schmit, 1982). In today's increasingly complex information environment, pilots maintain a running model of the situation to anticipate changes in the environment and then scan the display space to confirm or modify their mental model.

Most definitions of SA share common elements (e.g., Adams, et al., 1995; Endsley, 1988; Regal, et. al., 1988; Wickens, 1995). Wickens (1995) offers one of the best descriptions, describing SA in terms of a dynamic knowledge space, with emphasis on the aspects of SA that allow prediction and planning. Prediction and planning are a necessary consequence and benefit of "good SA." Wickens states that "SA is the continuous extraction of environmental information about a system or environment, the integration of this information with previous knowledge to form a coherent mental picture, and the use of that picture in directing further perception, anticipating, and responding to future events."

Wickens (1995) identified multiple factors affecting the maintenance of SA. Poor mental models of system complexities can cause control mode errors where pilots become confused about where they are in the information hierarchy of the glass cockpit. Other cognitive tasks such as electronic map scale, rotation, and dimensionality influence flight path guidance and path awareness. Additional critical pilot mental operations have been described as information integration, transformation, differentiation, and prediction (Jensen, 1978). While removing the pilot from the aircraft control loop through automation can maximize fuel economy and minimize deviation from a specified flight path, it also can result in some loss of SA on the part of the pilot. This loss of SA has a major impact on pilot effectiveness when the pilot must regain control of his aircraft. Wickens (1995) concluded that simultaneous automated support of flight guidance conflicts with ongoing maintenance of SA.

Complex information displays need to be designed that cluster functionally related data in ways that facilitate information acquisition across the display space. Modifications of cockpit environments, such as the addition of new capabilities, require a consideration of how the new functions will be integrated with the total information processing requirements. Integrated display designs will include decision aiding systems that keep the pilot in the control loop while providing adaptive or flexible automation that can adjust to the variable cognitive state of pilots (vigilance, boredom, fatigue, changing workload, and attention shifting). New skills and new training strategies will be required as pilots incorporate new functions into the cockpit.

Issues and Research

Information must be accessed in a way that maintains user control and enhances pilot performance under varying workload conditions. Information "hiding or layering" through the use of variable action buttons and multipurpose displays can impose additional cognitive and manual workload on pilots when it becomes necessary to retrieve data not currently displayed. Similarly, "notification by exception" and other approaches to decluttering (a decision not to display status information when all is well, for example)

can negatively impact planning ability - an important aspect of SA. Under periods of high stress or high workload," out-of-sight" can easily become "out-of-mind."

Critical to SA is information acquisition, fusion, and assessment at the right place, at the right time, with the proper use of graphical imagery. Sources of off-board information will increase in proportion to the increase in joint operations. Designers must improve the way pilots use and correlate information across multiple sources, some on-board and some off-board. Successful cockpit design depends on an accurate and predictive understanding of pilots' mental operations and dynamic information requirements.

What factors must be considered when simplifying complex displays intended to improve SA? For example, active management of graphical imagery could permit the pilot to selectively filter bird's-eye geographical displays using information overlays and decluttering symbologies. How much fidelity is required for SA? Too much fidelity could reduce SA by confusing pilots with superfluous or irrelevant information. Multi-sensory displays using augmented reality technologies, such as spatialized audio and see-through visual displays, can be used to optimize information loading and to cluster different functional classes of information. How can systems engineers exploit emerging technology and display capabilities to integrate information requirements, knowledge states, and multiple information delivery systems into a unitary cockpit design?

The following sections survey several emerging hardware-independent technologies that bear on these questions and that could help mitigate pilot information management requirements through cockpit display design.

1. Task Analysis and Cognitive Modeling of Pilot SA. Situational assessment can be modeled by thorough study of pilot task and information processing requirements, interface design and human-machine interaction. These task analyses should include a detailed assessment of what information is needed and when, as well as a comprehensive evaluation of how information is acquired and displayed across different workload conditions representing the full breadth of expected mission scenarios.

Workload analysis techniques that identify information bottlenecks in complex decision-making have been successfully applied to a wide spectrum of tasks from text editing to nuclear power plant decision-making (John and Kieras, 1994). Such interface and interaction analyses saved NYNEX millions of dollars by predicting seconds saved in telephone operator interface usage (Gray, John and Atwood, 1993). Recent advances in computational modeling of cognitive processes (see Meyer and Kieras, 1997a, b for review) have proven exceptionally accurate at predicting time-pressured decision-making performance with multiple tasks in the laboratory. The bottom line of these studies is that information processing limitations of interface design are due in large part to sensorimotor scheduling constraints in information acquisition rather than cognitive workload per se.

Future studies are needed to evaluate whether these laboratory studies will scale up to complex decision-making in the cockpit. If successful, such analytical techniques will reduce the valuable time required for actual usability studies with subject matter experts by providing human-in-the-loop specifications and constraints during the design process. In addition, such human-centered design approaches will provide insights into decision aiding and dynamic function allocation useful for task appropriate adaptive automation. The Man-Machine Integration Design and Analysis System (MIDAS) tool being developed at the NASA Ames Research Center (Corker, personal communication, 1997) is an attempt to integrate cognitive task analyses into more traditional ergonomic cockpit design for fixed-wing and rotorcraft platforms.

2. Three-Dimensional (3D) Visualization. Three-dimensional (3D) visualization technology is already in use and serves as one example of the utility of task and requirements definition. 3D representations of geographic information can improve the viewer's understanding of spatial relationships among objects within a volumetric context, a central component of SA. Conventional top-down two-dimensional views provide good estimates of relative location within a single plane but do not account for altitude or topographic contours. A static 3D view from an oblique or horizontal angle provides monocular cues including relative scale (e.g., vanishing point), relative position via occlusion and near field stereoscopic depth cues from retinal disparities. Dynamic interaction with the 3D view provides detailed depth information through motion parallax, the differential movement of objects as a function of depth. Three-dimensional images as a form of avionic cockpit display represent a potential benefit to pilot/aircraft performance for selected applications (Parrish and Williams, 1990; Parrish and Williams, 1992).

Despite the success of some attempts to incorporate 3D displays into cockpit systems, other research has shown that pilot tasks and displays interact in complex ways. Wickens (1995) summarizes a set of studies comparing 2D and 3D displays to support SA and has found there is no consistent advantage of one format over another, but rather the effectiveness depends on what particular subtask is being evaluated. For example, 3D displays have been shown to improve flight path control through the use of "tunnel-in-the-sky" volumetric imagery (Wickens and Prevett, 1995). Related preliminary research in the visualization of volumetric air traffic control information (Ballas, 1997) supports the notion that 3D is better suited for perceiving complex multidimensional relationships, tracking, and predicting movement in 3D space. Two dimensional representations are better suited for precision judgments on dimensions such as speed, heading, and vectoring. Analyses of specific information processing tasks are needed to identify when 3D versus 2D representations are more effective.

3. Decision aiding. Decision aids are another example of a cognitive intervention that becomes possible when a thorough analysis of information requirements has been done. Decision aids take many forms, but can generally be considered as comprising a set of techniques for gathering raw data from disparate sources and transforming it into useful information by imposing some increased structure, relevance, or richness to the data. In some decision aiding systems, recommended courses of action are given. In an effective decision aiding system, the human must remain in the decision process or control loop (Mitchell, 1993), in contrast to adaptive automation systems (see discussion below) in which an intelligent computer generally acts with more autonomy. Decision aiding implemented without explanation can be at odds with the maintenance and growth of cognitive skills in humans who have historically been included in the system to compensate for the limitation of the automation. Too much explanation has a great potential to generate increased workload, so some balance between cost and benefit must be achieved. Decision aiding in the form of information structuring and organization, has been shown to provide performance benefits to operators of other complex Navy systems, however (Hutchins, 1995).

Data fusion, or the integration of individual data streams into a more unified representation, is frequently considered as an example of decision aiding in that fusion eliminates the need for one or several cognitive steps by the pilot to get the information needed for action. The Naval Air Warfare Center, Aircraft Division (NAWC AD), has started a new ONR-funded program, “Effective Information Fusion for HMD Technologies,” to develop and evaluate HMD image-based concepts for application in tactical air missions. Navy researchers are teamed with Boeing Industries who is responsible for incorporating image-based concepts into a full mission flight simulation and conducting empirical evaluations.

4. Adaptive Automation. Adaptive automation of cockpit displays adjusts information content and format to fit the evolving information requirements of an ongoing mission and can be used to tailor displays to fit changing information requirements in real time. Requirements changes can be driven by a wide range of triggers including differences in pilot preferences and strengths, differing task demands over the course of a mission, or pilots’ physiological state. Changes in automation levels are frequently determined through a sensing mechanism incorporated into the onboard computing capability.

Cognitive task analysis can provide candidate tasks for off-loading through automation. Technologies that could be used in adaptive automation include dynamic information filtering, structuring, and intelligent agents as part-time or limited scope associates to fuse data and manage or filter information. Another ONR-funded program (NAWC AD), “Smart Cockpit Controller,” will design, develop, and demonstrate a controller to manage the functions of key cockpit technologies. The US Navy will define requirements, integrate key technologies/controller into F/A-18 simulator, and pilot-in-the-loop evaluations. The cockpit controller will be fabricated (based on US Navy input) for integration into the F/A-18 crewstation simulator by Boeing.

5. Pilot workload modeling/monitoring. As with the other approaches listed above, task analyses and cognitive modeling techniques are needed to build workload monitoring and management systems. This technology supports the assessment of the human-machine system to determine when to invoke assistive interventions. A variety of techniques range from physiological sensors (e.g., heart-rate monitoring, eye-tracking, electroencephalographic) (Van Orden, 1997) to computational analyses of cognitive workload - an extremely ambitious undertaking. ONR has recently provided funding at NAWC AD to continue development of in-flight physiological monitoring of tactical aircrews. The objective is to develop a real-time monitoring and feedback capability to restore and improve aircrew response to

stressors (e.g., acceleration, heat, altitude, fatigue, information overload). Advances are currently being made in alertness monitoring through eye and EEG monitoring at the Naval Health Research Center (NHRC) San Diego during low and high workload tasks over relatively long mission durations.

Discussion

Individually, none of the approaches listed here would be sufficient to manage the cockpit information center of the future. Any successful design will have to be carefully integrated with the hardware and human systems available. Thoughtful systems analysis is needed to determine the optimal allocation and balance of tasks between the hardware, software, and "liveware" components of the system.

As an example of the complexity of this problem, significant challenges exist when restricted to decisions about task allocation within the hardware system alone. Adam (1994) offers a candidate scheme for functional allocation between Head-Mounted Displays (HMDs) and a larger, "Big Picture -- Panoramic Cockpit" display for the 21st century. In this future design concept, the global situation would be presented on a relatively large (300 to 400 sq. in.) display, and would substitute, as least in part, for the display space provided by current head up displays (HUDs). The success of this concept would demand an effective tradeoff of functions currently displayed or performed via the HUD being accomplished by a much more extensive HMD capability. An HMD or Head Mounted System (HMS) would address the tactical portion of the SA problem (Adam, 1994).

To some, it is intuitively satisfying to imagine reduced workload and increased SA or "big picture" view with a larger, more integrated, display which fuses important information on the pilots' behalf. To others, the value of overlearned association of specific information with specific cockpit locations cannot be underestimated (Arvai, 1993). Arvai suggests that additional display area is desirable to allow immediate access to the information needed to maintain SA. The efficacy of competing approaches must be verified through analysis and testing that considers tradeoffs between design decisions, training intervention, pilot workload (task analysis), performance, and pilot preference and aviation tradition.

These research areas, listed briefly here, share an approach centered upon the human performance aspects of display systems design, where the largest gaps in our knowledge occur and the greatest potential for gains exist. To realize that potential, additional research in the application of cognitive technologies to complex systems is needed. There is growing recognition of the potential impact of cognitive technologies on future cockpit display systems among researchers and funding agencies, alike.

Current technologies including HUDs, HMDs, stereoscopic 3D display formats, and graphical summaries of aircraft performance information all offer the potential for significantly improved SA and piloting efficiency. The key to successful utilization of new technologies and approaches will be the effective integration of multiple displays with the information requirements of the decision-makers. This balance will provide the pilot with a concise picture of aircraft performance and environmental information in order to maintain control and make rapid adjustments to evolving conditions.

At the same time, overly zealous attempts to free the pilot from information overload through advanced display technologies could be detrimental to performance. In one recent study, difficulties in using cockpit automation were due to poor mental models of how automated systems work (Sarter and Woods, 1993). Mental effort spent attempting to understand the automated system only add to mental workload. Training in the design of automated pilot assistance systems is required for the pilot to understand and predict the hidden behavior of automated systems, creating a stronger partnership between pilot and system.

Emerging display technologies and presentation concepts must be applied in a top-down systems engineering approach that incorporates evaluation of new designs and consideration of the effects of modified designs on the generation and retention of SA. Task analysis and cognitive modeling techniques hold great promise for efficient integrated design that takes into account the information processing requirements of the pilot and air crew.

Simulators and virtual prototyping can be used as systems engineering tools to support avionics systems design. By using the same simulator for design development as the one that will eventually be used to run the training simulations, new design concepts can be integrated in iterative fashion. Moore and Moore (1985) advocate a structured, top-down procedure to describe displays, control, and operation. The approach would facilitate the implementation and evaluation of new design concepts, displays, information

management approaches, controls, and subsystems such as HMDs and voice synthesis and recognition systems.

General purpose systems engineering tools can be used to evaluate operator tasks, and manage task allocation between pilot and system. Task management should also include guidance for decision support or adaptive automation in the context of display clutter and information management. However, none of these design and evaluation techniques or cognitive interventions obviate the critical need for flight testing.

Conclusions

Technology alone is no panacea - while offering designers unlimited horizons, it can also overwhelm pilots with an ocean of data. The needs and capabilities of pilots and aircrews must be analyzed and accounted for. Effective systems engineering requires task analyses, function allocation, and predictive information processing models to facilitate information management and maximize SA in the cockpit. Designers of automated systems must carefully consider how the user interacts with the entire information display available to him, not just individual display systems. Cognitive modeling and task analytic tools will be key to improving SA through human-centered design.

References

Adam, E. C. "Head-up displays versus helmet-mounted displays: The Issues." Proceedings of the SPIE, Vol. 2465, pp. 175-183, Bellingham, WA: Society of Photo-Optical Instrumentation Engineers, April, 1994.

Adams, M. J., Tenney, Y. J., and Pew, R. W. State-of-the-art report: Strategic workload and the cognitive management of advanced multi-task systems (CSERIAC 91-6). Wright-Paterson Air Force Base, OH: Crew Systems Ergonomics Information Analysis Center, 1991.

AGARD 20/20 Report and Conference on Future Aerospace Technology in the Service Alliance. AGARD Advisory Report 360, Vol. 1. Advisory Group for Aerospace Research and Development. Palaiseau, France, 14-16 April, 1997.

Aronson, R. B. Cockpit of the future. Machine Design, pp. 455-459, 10 December, 1982.

Arvai, Joseph. V-22 cockpit management system analysis team report. Prepared for Naval Air Systems Command (AIR-5115X), 5 August 1993.

Ballas, J., Naval Research Laboratory, Washington, DC. Personal Communication, 1997.

Bost, J. R., Mellis, J. G. and Dent, P.A., "Is the Navy serious about reducing manning on its ships?" Presented at the Association of Scientists and Engineers of NAVSEA, 13 April, 1994.

Corker, K., Personal Communication, 27 January, 1997.

Endsley, M. R. Design and evaluation for situational awareness enhancement. In Proceedings of the Human Factors Society 32nd Annual Meeting, pp. 97-101, Santa Monica, CA: Human Factors and Ergonomics Society, 1988.

Gray, W. D., John, B. E., and Atwood, M. E. Project Ernestine: Validating a GOMS analysis for predicting and explaining real-world task performance. Human Computer Interaction, pp. 237-309, 1993.

Hutchins, S. and Rummel, K. A decision support system for tactical decision making under stress. In proceedings of the First International Symposium of Command and Control Research and Technology, pp. 163-173, 1995.

Jensen, R. S. The effects of prediction, quickening, frequency separation, and percent of pursuit in perspective displays for low visibility landing. In Proceedings of the Human Factors Society 22nd Annual Meeting, Santa Monica, Calif., Human Factors Society, Inc., pp. 208-212, 1978.

John, B. E. and Kieras, D. E. The GOMS family of analysis techniques: Tools for design and evaluation. Carnegie Mellon University Technical Report CMU-CS-94-181. 24 August 1994.

McGee, J. P. An advanced subsystem status monitor designed to reduce crew workload. Journal of the American Helicopter Society, pp. 10-15, January, 1982.

Meyer, D. E. and Kieras, D. E. A computational theory of executive cognitive processes and human multiple-task performance: Part 1. basic mechanisms. Psychological Review, in press, 1997.

Meyer, D. E. and Kieras, D. E. A computational theory of executive cognitive processes and human multiple-task performance: Part 2. accounts of psychological refractory-period phenomena. Psychological Review, in press, 1997.
Mitchell, C. M. Cognitive Engineering Models: A prerequisite to the design of human-computer interaction in complex dynamic systems. In Human Computer Interface Design: Success Cases, Emerging Methods, and Real World Context. Ed. P. Polson, New York: Morgan Kaufman, 1993.
Moore, D. and Moore, C. Pilot/vehicle interface design system. In National Aerospace and Electronics (NAECON); The National Aerospace and Electronics Conference, Dayton, OH, Vol. 2, pp. 1556-1563, 1985.
Parrish, R. and Williams, S. Stereopsis cueing effects on hover in turbulence performance in a simulated rotorcraft. NASA Technical Paper No. 2980, 1990.
Parrish, R. and Williams, S. Benefits, limitations, and guidelines for application of stereo 3D display technology to the cockpit environment. Paper presented at the NATO AGARD; Advanced Aircraft Interfaces, The Machine Side of the Man-Machine Interface Conference, Madrid, Spain, 1992.
Rockwell, T. H., Giffin, W. C., and Romer, D. J. Combining destination diversion decisions and critical inflight event diagnosis in computer aided testing of pilots. Proceedings of the 2nd Symposium on Aviation Psychology, 343- 351, 1983.
Roe, G. Electronic displays for tactical aircraft. Electronics and Power, The Journal of the Institution of Electrical Engineers, Vol. 28, No. 4, pp. 322-326, April 1982.
Regal, D. M., Rogers, W. H., and Boucek, G. P. Situational awareness in the commercial flight deck: definition, measurement, and enhancement. In Proceedings of the Seventh Aerospace Behavioral Technology Conference and Exposition, pp. 65-69, Warrendale, PA: Society of Automotive Engineers, 1988.
Sarter, N. B. and Woods, D. D. Cognitive engineering in aerospace application: pilot interaction with cockpit automation. NASA Contractor Report 177617. Ames Research Center August, 1993.
Schmit, V. P. Factors affecting the allocation of attention and performance in crossmonitoring flight information. displays. In AGARD Advanced Avionics and the Military Aircraft Man/Machine Interface, pp. 20.1-20.6, 1982.
Van Orden, K. Naval Health Research Center, San Diego; Personal Communication. 1997.
Wickens, C. D. "Situation awareness: Impact of automation and display technology." AGARD Conference Proceedings No. 575: Situation Awareness. Limitations and Enhancement in the Aviation Environment, Neuilly Sur Seine, France, pp. K2-1/K2-13, 1995.
Wickens, C and Prevett, T. "Exploring the dimensions of egocentricity in aircraft navigation displays: Influences on local guidance and global situation awareness." Journal of Experimental Psychology: Applied, Vol. 1, pp. 110-135, 1995.

*Currently on detail to the Office of Naval Research, Arlington, Virginia.

This page intentionally blank.

MULTI-SENSOR DATA FUSION AND SITUATIONAL AWARENESS IN THE TACTICAL COMMAND & CONTROL AIRCRAFT

Robert M. Chamberlain
CHI Systems, Inc.
2115 Jarrod Place
Smyrna, GA 30080

Floyd A. Glenn
CHI Systems, Inc.
716 N. Bethlehem Pike
Lower Gwynedd, PA 19002

Jay Wolf
CHI Systems, Inc.
716 N. Bethlehem Pike
Lower Gwynedd, PA 19002

There is a significant amount of important work going on currently in the area labeled "Situational Awareness," yet there does not seem to be any single definition of what exactly Situational Awareness is, and this lack of a definition has led to furious pockets of platform-specific situational awareness development activity, but any efforts that might link these platforms together seem to be missing. This linking together of information from the various available platforms, and the fusing together of data provided from these platforms into a useable "big picture" that can be distributed is a key to the future success of the tactical command & control platform.

Situational Awareness (SA) will mean different things to different people, depending on the environment in which they happen to be working. SA for a single seat attack pilot will be significantly different than SA for the Mission Commander in the E-2C, and will be significantly different than SA for the Joint Forward Area Control Center (JFACC), or Joint Task Force (JTF) Commanders. There are some common SA threads that flow through all levels, from the single seat attack pilot to the JTF Commander. Generally, current SA development work can be summed up as trying to answer the following questions:

> ***What is going on around me?***
> ***How will it affect my ability to accomplish my mission?***
> ***What, if anything, can I do about it?***

The bulk of the SA development work is directed at answering these questions on a platform-specific basis. The efforts in helmet-mounted displays, smart cockpits, adaptive automation, information display, and many other areas will make the pilot or aircrew member much more effective in employing their individual platforms and weapon systems. The important question that the current SA development work does not address well is:

> ***How do I pass the information I have along to the next higher level of decision making authority in a timely manner, and in a useable format?***

Current efforts seem to concentrate heavily on adding more sensors to platforms with a goal of making that platform more effective. Thus, we may have a tactical platform that has radar, Interrogation Friend or Foe (IFF), Infrared Search & Track (IRST), Synthetic or Inverse Synthetic Aperture Radar (SAR/ISAR), Non-Cooperative Identification (NCID) systems, ESM, low light TV, and other sensors that give the pilot and aircrew member(s) a much better capability to answer the first set of questions. The SA development work is concentrating on how to present the data gleaned from all of those sensors so as not to overwhelm individual aircrew(s), however, the methodology for passing this information up to the next level of decision-making authority is radio (plain or secure voice), Link 4a, Link 11, or JTIDS. This was the "70's" solution to the problem. SA development work must now address the issue of how to tie sensor information from a variety of different platforms together, present a "big picture" in a format that is easily understandable and useable, and then to pass the information on through the established chain of command so that timely evaluation and decision-making can occur.

Multi-sensor Data Fusion.

Multi-sensor data fusion is another term that means different things to different people. There are a variety of sensors in the Navy's current tactical command & control aircraft (the E-2C), including radar, IFF,

ESM, and several modes of communications that include voice and data link. SATCOM data capability, Cooperative Engagement (CEC), and IRST are planned additions to the aircraft weapon system in the near future as well. The problem is that currently the *operator* is the multi-sensor data fuser. The operator must get out of the business of being the multi-sensor data collector, and make the system do that. The operator must be a manager of information and a decision-maker, which means that the on-board system must be capable of putting together all of the data provided from the variety of on-board and off-board sensor systems, and presenting the fused "data" as information in a useable and understandable format so that the operator can make timely and rational decisions based on that information.

Currently in the E-2C, the system attempts to automatically correlate radar and IFF returns, and "fuse" the information so that dual tracks are not presented to the operator. The ESM system (generally referred to as Passive Detection System (PDS)) is a stand alone sensor system that will present data on the same operator display, but any attempt to correlate or fuse ESM information with that of radar and IFF is a manual effort undertaken by the operator. Thus, while there may be a wealth of information available from the ESM system, the process of trying to correlate this data with the data provided from the radar, IFF, and communications systems is so operator intensive that the operator can quickly lose SA (*what is going on around me, how will it affect my mission, and what can I do about it?).* Clearly, an important step in SA development for the E-2C is the fusing of ESM data with other sensor data, and providing the "fused" information to the operator. This step will greatly enhance the SA of the operator.
The need for multi-sensor data fusion of current sensor systems in the E-2C to improve SA is obvious. The growth path for the aircraft indicates the addition of other sensor and data systems such as CEC, IRST, and SATCOM data so that the aircraft will have the capability to address new missions (theater ballistic missile warfare, weapons guidance for the Arsenal Ship or Aegis ships, UAV control, and others). The E-2C operator will be tasked with completing these new missions in the near future, but not at the expense of the existing missions, as these will remain. To adequately address the addition of sensors to the platform to accomplish these new missions requires fusion of data from new sensor systems with existing sensor and communications systems. Thus, if an IRST system is added to the platform, the approach should not be one of determining where an additional display could be added to the weapons system to present data to the operator, but how to take the data output of the IRST system and integrate it with data from the other systems, and provide information to the operator on the main display that allows the operator to evaluate and make decisions.

On-board & Off board Sensor Fusion

The need for fusion of data provided by on-board sensor systems on the E-2C is no different than for any other tactical aircraft. The key issue is providing a clear picture to the pilot, or operator, so that an evaluation can be made (*what is going on around me?*). Unlike the majority of tactical aircraft, the E-2C has off-board data and sensor information coming in from a variety of other platforms. Currently, the data coming into the aircraft is typically in the form of voice communications (either plain voice or secure voice) or data link (Link 4a, Link 11, and JTIDS). The addition of SATCOM data capability will dramatically increase the sources and volume of data coming into the aircraft. SATCOM data will provide the opportunity to significantly increase the SA of the E-2C NFO by providing a much more detailed set of information on what is going on around the platform, and by significantly increasing the range of SA. No longer will the E-2C NFO be limited by the range of the on-board sensors. The E-2C NFO can receive a theater-wide picture, rather than a 500 mile glimpse of the tactical situation. The challenge for increasing SA will be the ability to fuse data from this wide variety of sources, and provide reliable, useful information to the NFO. The displayed information will indicate what is going on around the platform, so that the pilot can evaluate the implications, make recommendations (rather than just report) to the next higher level of authority, and make decisions on what to do about it.

Target Identification

The critical aspect of SA (*what's going on around me?)* is the ability to identify, to some level of assurance, the targets in the tactical environment. Here again, significant technical strides have been taken

in a number of different sensor systems to provide high quality target identification. ESM systems, Jet Engine Modulation (JEM) systems, SAR/ISAR radars, FLIR's, and other sensor systems all have the capability to provide some form of target identification. We believe that we no longer have the time to wait for a visual identification of a target by a friendly aircraft, and we have proven (in the friendly helicopter shoot-down in northern Iraq), that even a visual identification is not always reliable. A significant problem in the tactical command & control aircraft is that there currently is no means of exchanging and correlating identification information provided by this wide variety of target ID sensor systems, other than voice communications. We must establish a means of receiving all of this data, and integrate a multi-sensor data fusion capability that will absorb this wide variety of sensor data, perform an evaluation based on a set of established criteria, and display it to the operator with a designated level of confidence that the target actually is what the system says it is. This will give the operator the flexibility to either believe or not believe the system evaluation, make decisions, and take appropriate action.

Platform Interaction

The majority of multi-sensor data fusion work for increasing SA is oriented toward individual tactical aircraft. The E-2C viewpoint is that multi-platform interaction is required to significantly increase SA. The U.S. Military has a variety of different platforms with extremely capable sensor systems, but no good methodology for systematically exchanging and correlating (in real time) the information that these platforms provide. Platform examples include the F/A-18 with a JEM system, an S-3 with a FLIR, a P-3 with a SAR radar, and non-Navy platforms such as the JSTARS, EC-135's, AWACS, Hawk missile batteries, and a wide variety of other sensor platforms. All of these platforms are very capable in their assigned missions, and have sensor systems designed to provide the individual pilots and air crew members with SA and the ability to accomplish their respective missions.

The missing piece is the ability for these platforms to interface with each other from a data exchange standpoint, providing a collective picture that is more comprehensive than that available on the individual platform. It is this collective picture that will provide enhanced SA for the E-2C aircrew, and other command & control platforms. We are currently accomplishing this using voice communications and rudimentary data links. A significant development in multi-sensor data fusion, in conjunction with a development effort in multi-platform data exchange, has the potential to enhance SA from the single seat tactical cockpit all the way through to the battle staff.

Jointness Issues.

The exchange of data between platforms, and presentation of that data to aircrew in a clear, comprehensive, and useable format is critical to our ability to truly enhance SA in the overall warfighting structure. We cannot afford to limit our efforts to "Navy only" platforms. The other services have platforms with excellent sensor systems that can compliment Navy platforms when used together to accomplish a mission. Typically, the services do not spend enough time training together to understand the capabilities and limitations of their respective platforms, and the contributions that could be made to joint operations. A part of this general problem is that no methodology, data exchange systems, hardware, software, or display concepts exist to address this problem, and this area does not appear to be the focus for future SA development. Again, most of the SA development effort is going into single platform enhancements, when greater strides could be taken, and much more accomplished, if the focus were shifted to this issue of multi-service, multi-platform information sharing and fusion.

A good example of this is the JSTARS aircraft, an Air Force platform with a very capable, specialized sensor system. A logical assumption would be that the output of this sensor system would be passed to command & control aircraft such as AWACS or E-2C's for fusion with other data. The composite picture provided would be invaluable for strike aircraft, close air support aircraft, and ground commanders. Discussions with JSTARS aircrew indicate that there is no data exchange interface built into the system to coordinate with, or direct tactical forces. Thus, while the system is excellent, there is no means for the

aircrew to provide the information gleaned from this system to a command & control network in anything close to real time. This seems to be an area with real potential for further SA development.

Conclusion

The current efforts in SA development are laudable, and for tactical aircrew, represent some excellent efforts. The orientation, and emphasis has been primarily on increased SA at the single platform level directed at survivability enhancement. This is necessary in today's "accomplish the mission with no losses" environment. We must now take the next step in SA development, and address how SA can be enhanced at the command & control level, and how fused information taken from all of our available platforms can be placed on displays in cockpits, on ships, and at ground stations. This level of SA enhancement will dramatically increase our warfighting ability.